

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2004, 15:12:04 ; Search time 8805 Seconds
(without alignments)
11326.783 Million cell updates/sec

Title: US-10-043-715-1_COPY_157700_160000
Perfect score: 2301
Sequence: 1 ctgaatgatgtctcccaagt.....tagctccacggagagccct 2301

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.scs.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pin.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2301	100.0	186510	9	HS451B15	Z98050 Human DNA s
2	2292.6	99.6	109047	2	HSDB19F5	AL078592 Homo sapi
3	2273.2	98.8	10006	9	AY434104	AY434104 Homo sapi
4	2241.6	97.4	12461	9	HUMEDN1B	J05008 Homo sapien
5	2230	96.9	12460	6	A98526	A98526 Sequence 1
6	2230	96.9	12460	6	BD080748	BD080748 Antisense
7	662.4	28.8	684	9	HUMETN3	M25379 Human endot
8	411.4	17.9	1425	9	AC036851	BC036851 Homo sapi
9	263.2	11.4	259474	2	AC095692	AC095692 Rattus no
10	241	10.5	1330	9	HUMETN4	M25380 Human endot
11	239	10.4	237394	2	AC131424	AC131424 Rattus no
12	166	7.2	266	6	AF130760	AF130760 Equus cab
13	157.4	6.8	636	6	BD094780	BD094780 The cell
14	157.4	6.8	636	6	BD096322	BD096322 Cells cap
15	157.4	6.8	1175	6	E02620	E02620 DNA encodin
16	157.4	6.8	1175	6	AX587988	AX587988 Sequence
17	157.4	6.8	1251	9	S56805	S56805 preproendot
18	157.4	6.8	1251	9	BC009720	BC009720 Homo sapi
19	157.4	6.8	1334	9	MFU20579	U20579 Macaca fasc
20	150.2	6.5	230	9	D7083952	D70840 Mus musculu
21	125	5.4	139	10	AB081657	AB081657 Mus muscu
22	123.6	5.4	98530	2	AB113088	AB113088 Mus muscu
23	112.4	4.9	1936	4	AB115087	AB115087 Canis fam
24	111.2	4.8	447	10	S82654	S82654 preproendot
25	107.2	4.7	949	4	BTENDOTH	X52942 Bovine mRNA
26	107.2	4.7	1762	4	BTENDOTH	X52740 Bovine mRNA
27	107.2	4.7	2023	4	S37093	S37093 preproendot
28	103	4.5	1385	10	RATET1	M64711 Rat endothe
29	101.4	4.4	695	6	AX525996	AX525996 Sequence
30	100.8	4.4	749	6	E02619	E02619 DNA encodin
31	100.8	4.4	1838	4	SSET	X07383 Porcine mRN
32	99.2	4.3	626	10	AB081657	AB081657 Mus muscu
33	99.2	4.3	1770	10	MMU35233	U35233 Mus musculu
34	99.2	4.3	2153	10	BC029547	BC029547 Mus muscu
35	99	4.3	1225	11	AB031407	AB031407 Sus scrof
36	97.6	4.2	1143	4	AF329468	AF329468 Ovis arie
37	96.8	4.2	1639	4	OCET1	X59931 O. cuniculus
38	96	4.2	1586	10	D43775	D43775 Mouse mRNA
39	93.4	4.1	319	10	AF122903	AF122903 Rattus no
40	88.2	3.8	288	4	AF333433	AF333433 Canis fam
41	79	3.4	1231	10	D7083954	D70842 Mus musculu
42	73	3.2	236	4	AF320770	AF320770 Felis cat
43	70	3.0	154	10	D7083953	D70841 Mus musculu
44	66.6	2.9	7218	6	I66494	I66494 Sequence 14
45	58	2.5	1141	6	AX083744	AX083744 Sequence

ALIGNMENTS

RESULT 1
HS451B15
LOCUS
DEFINITION
Human DNA sequence from clone RP3-451B15 on chromosome 6p24
Contains the 3' end of the HIVEP1 gene for human immunodeficiency
virus type 1 enhancer-binding protein 1, the EDN1 gene for
endothelin 1 (ET1) and a SMR3 suppressor of mlf two 3 homolog 1
(yeast) (SMT3H1) pseudogene, complete sequence.

ACCESSION

Z98050

Z98050.1 GI:2791272

VERSION

HTG; EDN1; endothelin; ET1; HIVEP1; SMT3H1.

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

REFERENCE AUTHORS TITLE JOURNAL

1 (bases 1 to 186510)
Tubby,B.
Direct Submission
Submitted (05-JUN-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SN, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Jan 18, 1998 this sequence version replaced gi:2558561.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquerry@sanger.ac.uk

COMMENT

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the rare
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information
on the WormPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr6>
RP3-451B15 is from the library RPCI-3 constructed by the group of
Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
VECTOR: pCYPAC2

FEATURES source

This sequence is the entire insert of clone RP3-451B15.

Location/Qualifiers
1..186510
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="RZPD:RPCIP704B15451"
/db_xref="taxon:9606"
/chromosome="6"
/map="p24"
/clone="RP3-451B15"
/clone_lib="RPCI-3"

repeat_region
8..17
/note="2.0 copies 5 mer CCTCT 20% conserved"
repeat_region
25..44
/note="2.0 copies 10 mer GGTTCCTATC 31% conserved"
repeat_region
253..325
/note="L2 repeat: matches 2985..3059 of consensus"
repeat_region
320..331
/note="2.0 copies 6 mer TTTTCC 24% conserved"
repeat_region
767..781
/note="15.0 copies 1 mer A 21% conserved"
repeat_region
1044..1060
/note="2.1 copies 8 mer GGAAGACT 34% conserved"
repeat_region
1432..1448
/note="2.4 copies 7 mer TTTTACC 25% conserved"
repeat_region
1499..1509
/note="11.0 copies 1 mer T 22% conserved"
repeat_region
1711..1728
/note="1.9 copies 10 mer TGTGCTTTC 38% conserved"
repeat_region
2198..2207
/note="2.5 copies 4 mer TTTG 20% conserved"
repeat_region
2415..3238
/note="L2 repeat: matches 2058..2971 of consensus"

repeat_region
2500..2511
/note="2.4 copies 5 mer TCTCT 24% conserved"
repeat_region
2874..2885
/note="2.4 copies 5 mer CACCC 24% conserved"
repeat_region
3247..3257
/note="2.2 copies 5 mer ATTCC 22% conserved"
repeat_region
3255..3417
/note="L1PA6 repeat: matches 11..6154 of consensus"
repeat_region
3374..3403
/note="2.0 copies 15 mer ACAGAGGAGGCC 42% conserved"
repeat_region
4042..4074
/note="1.9 copies 17 mer CAATATTGCGTCTG 57% conserved"
repeat_region
4173..4187
/note="3.8 copies 4 mer AAAC 21% conserved"
repeat_region
4196..4209
/note="2.3 copies 6 mer CATCAA 28% conserved"
repeat_region
4301..4311
/note="2.2 copies 5 mer ACCAG 22% conserved"
repeat_region
4341..4351
/note="2.2 copies 5 mer CTCTT 22% conserved"
repeat_region
4686..4696
/note="2.2 copies 5 mer GAGAA 22% conserved"
repeat_region
5112..5122
/note="2.2 copies 5 mer AAAGA 22% conserved"
repeat_region
5578..5606
/note="29.0 copies 1 mer A 56% conserved"
repeat_region
6181..6191
/note="2.2 copies 5 mer AATGA 22% conserved"
repeat_region
6535..6554
/note="6.7 copies 3 mer AAG 24% conserved"
repeat_region
7034..7044
/note="2.2 copies 5 mer CAAAC 22% conserved"
repeat_region
8046..8057
/note="2.0 copies 6 mer TGGTAC 24% conserved"
repeat_region
8081..8092
/note="2.4 copies 5 mer GAACA 24% conserved"
repeat_region
8863..8874
/note="2.0 copies 6 mer CTAGAA 24% conserved"
repeat_region
9402..9419
/note="18.0 copies 1 mer A 36% conserved"
repeat_region
9420..9776
/note="L2 repeat: matches 2962..3309 of consensus"
repeat_region
9508..9537
/note="15.0 copies 2 mer AT 26% conserved"
repeat_region
9618..9627
/note="3.3 copies 3 mer ACC 20% conserved"
repeat_region
9838..9847
/note="2.5 copies 4 mer CTGC 20% conserved"
repeat_region
10047..10336
/note="AluSp repeat: matches 1..290 of consensus"
repeat_region
10335..10352
/note="6.0 copies 3 mer AAT 27% conserved"
repeat_region
10493..10519
/note="2.5 copies 11 mer TTATTATTGG 36% conserved"
repeat_region
10508..10521
/note="2.0 copies 7 mer TTATTGG 28% conserved"
repeat_region
complement(10584..10652)
/note="L3b repeat: matches 1422..1493 of consensus"
repeat_region
10605..10615
/note="2.2 copies 5 mer TCTGT 22% conserved"
repeat_region
10661..10672
/note="12.0 copies 1 mer T 24% conserved"
repeat_region
10828..10843
/note="4.0 copies 4 mer GAAG 23% conserved"
repeat_region
10831..10841
/note="2.2 copies 5 mer GGAAG 22% conserved"
repeat_region
11401..11427
/note="6.8 copies 4 mer GTTT 40% conserved"
repeat_region
11425..11436
/note="2.0 copies 6 mer TTGCTG 24% conserved"
repeat_region
11521..11533
/note="2.2 copies 6 mer TAATAC 26% conserved"
repeat_region
11691..11706

```
/note="2.0 copies 8 mer ATGTATAA 32% conserved"
13249..13265
/note="2.1 copies 8 mer GTTTATAG 25% conserved"
13345..13354
/note="2.5 copies 4 mer TTCA 20% conserved"
13531..13542
/note="2.4 copies 5 mer TTGTT 24% conserved"
13739..13748
/note="2.0 copies 5 mer TGATG 20% conserved"
14533..14548
/note="2.3 copies 7 mer TAGGATT 23% conserved"
14673..14685
/note="6.5 copies 2 mer AT 26% conserved"
14887..14898
/note="2.0 copies 6 mer CGACTC 24% conserved"
15201..15210
/note="2.5 copies 4 mer ATTT 20% conserved"
/note="15548..15764) complement(15548..15764)
15779..16079
/note="AluX repeat: matches 20..312 of consensus"
16050..16064
/note="15.0 copies 1 mer A 21% conserved"
16066..16075
/note="3.3 copies 3 mer ACA 20% conserved"
16294..16306
/note="2.2 copies 6 mer GGGCA 26% conserved"
16739..16753
/note="3.0 copies 5 mer CTGA 30% conserved"
16919..16938
/note="2.2 copies 9 mer ATTTAAAT 31% conserved"
17193..17247
/note="MIR repeat: matches 32..85 of consensus"
17458..17479
/note="22.0 copies 1 mer A 44% conserved"
17752..17978
/note="AluSg repeat: matches 1..227 of consensus"
complement(17993..18461)
/note="L1M4a repeat: matches 6126..6606 of consensus"
18482..18708
/note="L1P8 repeat: matches 5948..6162 of consensus"

Query Match 100.0%; Score 2301; DB 9; Length 186510;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGAATGATGCTCCCAAGTGTATGATGAGCTCCTTGTGTCGCCAGTGGATAGTG 60
Db 157700 CTGAATGATGCTCCCAAGTGTATGATGAGCTCCTTGTGTCGCCAGTGGATAGTG 157759

Qy 61 TGTCCATGTGCTATTTAAAGACTATTAATACACTAATATAGTTTCTTCTCTCTTTGG 120
Db 157760 TGTCCATGTGCTATTTAAAGACTATTAATACACTAATATAGTTTCTTCTCTTTGG 157819

Qy 121 ATATAGCAGCTGTTTCGGTATGAGCTTGAAGCCCTAGTCCAGAGAGCCCTGGAGA 180
Db 157820 ATATAGCAGCTGTTTCGGTATGAGCTTGAAGCCCTAGTCCAGAGAGCCCTGGAGA 157879

Qy 181 ATTTACTTCCCAAGGCAACAGACCGTGAATAGATGCAATGTGCTAGCCAAAG 240
Db 157880 ATTTACTTCCCAAGGCAACAGACCGTGAATAGATGCAATGTGCTAGCCAAAG 157939

Qy 241 ACAGAAAGTGTGGAATTTTCCAGCAGGAAAGAACTCAGGTGAGCAACAACCTT 300
Db 157940 ACAGAAAGTGTGGAATTTTCCAGCAGGAAAGAACTCAGGTGAGCAACAACCTT 157999

Qy 301 TGTCTTCAATCAGTTTAAACAGCTCCCTGAACTCCTTCTATCATGTGCTTCCCTG 360
Db 158000 TGTCTTCAATCAGTTTAAACAGCTCCCTGAACTCCTTCTATCATGTGCTTCCCTG 158059

Qy 361 TTTTAGAGACTAACAGAGACTTGAAGTCAGGCTAAAGCTGAATATACATTCGCTGA 420
Db 158060 TTTTAGAGACTAACAGAGACTTGAAGTCAGGCTAAAGCTGAATATACATTCGCTGA 158119
```

```
Qy 421 AATGTTTTCTCTGTGTATTTTAAACAGGCTGAAGACATTTATCGAAGAACTGGAATAA 480
Db 158120 AATGTTTTCTCTGTGTATTTTAAACAGGCTGAAGACATTTATCGAAGAACTGGAATAA 158179

Qy 481 TCATAAGAAAGGAAAGACATGTTTCCAGCTTGGAAAAGTGTATTTATCAGCAGTTAGT 540
Db 158180 TCATAAGAAAGGAAAGACATGTTTCCAAAGCTTGGAAAAGTGTATTTATCAGCAGTTAGT 158239

Qy 541 GAGAGGAAGAAAATCAGAGAAAGTTTCAAGAGAACCTTAAGACAACACAGGTAAGAGGG 600
Db 158240 GAGAGGAAGAAAATCAGAGAAAGTTTCAAGAGAACCTTAAGACAACACAGGTAAGAGGG 158299

Qy 601 AAGCAAGAAAATTTAGGTAAAGGTTTCAAGAAACAATAGCCCCAGTCAGTGATGCCAG 660
Db 158300 AAGCAAGAAAATTTAGGTAAAGGTTTCAAGAAACAATAGCCCCAGTCAGTGATGCCAG 158359

Qy 661 CAGCTGTTCTCCAGCCCTTCTTCCGGGCAAGTGAAGACTTAGAAAACAGTAGACAG 720
Db 158360 CAGCTGTTCTCCAGCCCTTCTTCCGGGCAAGTGAAGACTTAGAAAACAGTAGACAG 158419

Qy 721 AGAGAGTCTATGCATCTATAGATTTAAAGAGCAAAAGATCCCTCTTAAATATTTCCA 780
Db 158420 AGAGAGTCTATGCATCTATAGATTTAAAGAGCAAAAGATCCCTCTTAAATATTTCCA 158479

Qy 781 TGAAGCTCTGGAATGCAACCCGATGCTCTGTACTTTTAGCACATACCATTTCATCTAC 840
Db 158480 TGAAGCTCTGGAATGCAACCCGATGCTCTGTACTTTTAGCACATACCATTTCATCTAC 158539

Qy 841 AGGTAGATTTCCCAACCAAAATATATCCAGAGATGCTTGTTCATTTGGTTATATACAG 900
Db 158540 AGGTAGATTTCCCAACCAAAATATATCCAGAGATGCTTGTTCATTTGGTTATATACAG 158599

Qy 901 CTTTGGCTCTCTGAGTCAATGTATTTACCACTTTCCTCAGAAATCGAAAATCATTTTG 960
Db 158600 CTTTGGCTCTCTGAGTCAATGTATTTACCACTTTCCTCAGAAATCGAAAATCATTTTG 158659

Qy 961 GGAGCGGACATTTAGAAAAGAAATCAAGTGCATGATTAATCAATTTCTTCAATAGTT 1020
Db 158660 GGAGCGGACATTTAGAAAAGAAATCAAGTGCATGATTAATCAATTTCTTCAATAGTT 158719

Qy 1021 GCAGTTATTTCAGATGCGCCAAAGGAAAATAAAGTCAATTAGATAGGTTTGGTGAATTTAG 1080
Db 158720 GCAGTTATTTCAGATGCGCCAAAGGAAAATAAAGTCAATTAGATAGGTTTGGTGAATTTAG 158779

Qy 1081 AACATGCTGTTTTTCAGGTTTATGGTCTTTTTTTTTTTTTTTTTTTTTTAAATAGGAAATG 1140
Db 158780 AACATGCTGTTTTTCAGGTTTATGGTCTTTTTTTTTTTTTTTTTTTTTTAAATAGGAAATG 158839

Qy 1141 TGTTCGTCGAGAGCCCAATGTCTATTTCCAAAAGCTCTCTCTTTTCTGCTCAGTCATGTG 1200
Db 158840 TGTTCGTCGAGAGCCCAATGTCTATTTCCAAAAGCTCTCTCTTTTCTGCTCAGTCATGTG 158899

Qy 1201 CTGGGACAGAGAGGATCTGGATTAGGCAACATCATAGAGTTGCTCTGAGCTGCTCTTT 1260
Db 158900 CTGGGACAGAGAGGATCTGGATTAGGCAACATCATAGAGTTGCTCTGAGCTGCTCTTT 158959

Qy 1261 GGTGATAACCCCTTCCAAATCCCTAACTTTTGGAAATTCACAAGCTCAAGAGGAGAAACCT 1320
Db 158960 GGTGATAACCCCTTCCAAATCCCTAACTTTTGGAAATTCACAAGCTCAAGAGGAGAAACCT 159019

Qy 1321 ACTCTCTGATCTACCAATGTTCTGCAATTTTCTATCATGTTCTATGGAACCTTCTCTTA 1380
Db 159020 ACTCTCTGATCTACCAATGTTCTGCAATTTTCTATCATGTTCTATGGAACCTTCTCTTA 159079

Qy 1381 GAAATCCAGTGCAGAGAGTTCTATGATTTAAAGTGTCTGAGCTCAGGCGAGGAGTCAAT 1440
Db 159080 GAAATCCAGTGCAGAGAGTTCTATGATTTAAAGTGTCTGAGCTCAGGCGAGGAGTCAAT 159139

Qy 1441 GAACTACTTCTGAGTTTATTTACTACTGATTTTGTGGGAGCCCTCAGCTATCGGTTCTTC 1500
Db 159140 GAACTACTTCTGAGTTTATTTACTACTGATTTTGTGGGAGCCCTCAGCTATCGGTTCTTC 159199
```

QY	1501	ACACCTGCTTATGAGAGTATCCATATTTATGGTCGACGGCCAGTAATGCTCCCCACGAGA	1560
Db	159200	ACACCTGCTTATGAGAGTATCCATATTTATGGTCGACGGCCAGTAATGCTCCCCACGAGA	159259
QY	1561	TCAGTTTCTGAACCTAAGCTGAATTTTATGGGTTTATTTATGCGCAACTATTAAATCA	1620
Db	159260	TCAGTTTCTGAACCTAAGCTGAATTTTATGGGTTTATTTATGCGCAACTATTAAATCA	159319
QY	1621	ACATTACAGTTCTTCCCTCTGTATTTCTCTGTAATAACATTTAGCGCTGCAAAAAA	1680
Db	159320	ACATTACAGTTCTTCCCTCTGTATTTCTCTGTAATAACATTTAGCGCTGCAAAAAA	159379
QY	1681	ATCTTTTAAAAATAATTTGCCATAAAGTATTTGCTCTGGGCTACTGTATGCTCTTTTC	1740
Db	159380	ATCTTTTAAAAATAATTTGCCATAAAGTATTTGCTCTGGGCTACTGTATGCTCTTTTC	159439
QY	1741	TTTTTCTCTCTTTTCAACTAAGTCACCGTCAATTTATTAAGATGCGCATTAATTCAAA	1800
Db	159440	TTTTTCTCTCTTTTCAACTAAGTCACCGTCAATTTATTAAGATGCGCATTAATTCAAA	159499
QY	1801	ACCTATGCTGAGTTCTTCAAGCGAGGTCACATAGTGATGAAGTTTGGATGGGGCTACG	1860
Db	159500	ACCTATGCTGAGTTCTTCAAGCGAGGTCACATAGTGATGAAGTTTGGATGGGGCTACG	159559
QY	1861	GAAGAAACCCAGAACAACTCTAGTTTATTTAAACCTGTATTTACTGCCACTTCCCCTTA	1920
Db	159560	GAAGAAACCCAGAACAACTCTAGTTTATTTAAACCTGTATTTACTGCCACTTCCCCTTA	159619
QY	1921	GACTTGACCATATGACCCCTCGCTCCCATTTCTAAGCATAGGGCAGGCTTTATTTTACA	1980
Db	159620	GACTTGACCATATGACCCCTCGCTCCCATTTCTAAGCATAGGGCAGGCTTTATTTTACA	159679
QY	1981	ATGGTAAATAGATATCATTGAGTTTATCAAAGATTTGGCGGGTGGTGAAGTTTCAC	2040
Db	159680	ATGGTAAATAGATATCATTGAGTTTATCAAAGATTTGGCGGGTGGTGAAGTTTCAC	159739
QY	2041	AACAGATTCAGGTTTGTGTCGAGATTCCTAAATTTTACATGTTTCTTTTGCCAAAGG	2100
Db	159740	AACAGATTCAGGTTTGTGTCGAGATTCCTAAATTTTACATGTTTCTTTTGCCAAAGG	159799
QY	2101	GTGATTTTTTAAATAACATTTGTTTCTCTATCTTCTTATTTAGTTCGGAGCCAT	2160
Db	159800	GTGATTTTTTAAATAACATTTGTTTCTCTATCTTCTTATTTAGTTCGGAGCCAT	159859
QY	2161	GAGAAACAGGTCACAAATCATCTTTTCATGATCCCAAGCTGAAAGCAAGCCCTCCAGAGA	2220
Db	159860	GAGAAACAGGTCACAAATCATCTTTTCATGATCCCAAGCTGAAAGCAAGCCCTCCAGAGA	159919
QY	2221	GGTTATGTGACCCCAACACCGAGACATTTGTGACAGACCTTGGGGCTGTCTGAAGCC	2280
Db	159920	GGTTATGTGACCCCAACACCGAGACATTTGTGACAGACCTTGGGGCTGTCTGAAGCC	159979
QY	2281	ATAGCTTCCACGAGAGGCCCT 2301	
Db	159980	ATAGCTTCCACGAGAGGCCCT 160000	
RESULT 2			
HSDJ19F5/c			
LOCUS	HSDJ19F5 109047 bp DNA linear HTG 10-JUL-2001		
DEFINITION	Homo sapiens chromosome 6 clone RP1-19F5 map p22.3-24.2, 3		
ACCESSION	AL078592		
VERSION	AL078592.6 GI:9931903		
KEYWORDS	HTG; HTGS PHASE1; HTGS_CANCELLED.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1		
TITLE	Griffiths, C.		
JOURNAL	Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire.		
COMMENT			
CB10 18A, UK. E-mail enquiries: humquery@sanger.ac.uk Clone			
requests: clonerequest@sanger.ac.uk			
On Aug 28, 2000 this sequence version replaced gi:9864399.			
----- Genome Center			
Center: Sanger Centre			
Center code: SC			
Web site: http://www.sanger.ac.uk			
Contact: humquery@sanger.ac.uk			
----- Project Information			
Center project name: dJ19F5			
----- Summary Statistics			
Assembly program: XGAP4; version 4.5			
Sequencing vector: M13; M7815; 55% of reads			
Sequencing vector: plasmid; L08752; 44% of reads			
Chemistry: Dye-terminator ET-amersham; 0% of reads Chemistry:			
Dye-terminator ABI; 31% of reads			
Chemistry: Dye-terminator Big Dye; 43% of reads			
Chemistry: Dye-primer-amersham; 24% of reads			
Consensus quality: 108264 bases at least Q40			
Consensus quality: 108412 bases at least Q30			
Consensus quality: 108614 bases at least Q20			
Insert size: 108847; sum-of-contigs			
Insert size: 123663; 8.0% error; agarose-fp			
Quality coverage: 9.30x in Q20 bases; sum-of-contigs Quality			
coverage: 8.80x in Q20 bases; agarose-fp			

* NOTE: This is a 'working draft' sequence. It currently			
* consists of 3 contigs. The true order of the pieces			
* is not known and their order in this sequence record is			
* arbitrary. Gaps between the contigs are represented as			
* runs of N, but the exact sizes of the gaps are unknown.			
* This record will be updated with the finished sequence			
* as soon as it is available and the accession number will			
* be preserved.			

1 80880: contig of 80880 bp in length			
80881 80980: gap of 100 bp			
80981 99461: contig of 18481 bp in length			
99462 99561: gap of 100 bp			
99562 109047: contig of 9486 bp in length.			

Location/Qualifiers			
1..109047			
/organism="Homo sapiens"			
/mol_type="genomic DNA"			
/db_xref="taxon:9606"			
/chromosome="6"			
/map="p22.3-24.2"			
/clone="RP1-19F5"			
/clone_lib="RPC1-1"			
1..80880			
/note="assembly_fragment:00652"			
80981..99461			
/note="assembly_fragment:00992"			
99562..109047			
/note="assembly_fragment:02363"			

misc_feature			
misc_feature			
misc_feature			

ORIGIN			
Query Match 99.6%; Score 2292.6; DB 2; Length 109047;			
Best Local Similarity 99.8%; Pred. No. 0;			
Matches 2295; Conservative 0; Mismatches 4; Indels 0; Gaps 0;			
QY	1	CTGAATATGCTCCCAAGTCTATGTATGATGAGCTCTTGTGTGCCAGTGAATAGGTG	60
Db	2415	CTGAATATGCTCCCAAGTCTATGTATGATGAGCTCTTGTGTGCCAGTGAATAGGTG	2356
QY	61	TGTCATGCTCATTTTAAAGACTATTATTAACATAATATAGTTCTTCTCTCTTGG	120
Db	2355	TGTCATGCTCATTTTAAAGACTATTATTAACATAATATAGTTCTTCTCTCTTGG	2296
QY	121	ATAATAGGACGCTTGTTCGTATGACTTGGAAAGCCCTAGGTCCCAAGAGAGCCTTGGAGA	180
Db	2295	ATAATAGGACGCTTGTTCGTATGACTTGGAAAGCCCTAGGTCCCAAGAGAGCCTTGGAGA	2236
QY	181	ATTACTTCCCAAGGCAACAGACCGGTGAATATAGTGCCTAGCCAAAAG	240

	Db	2235	ATTTACTTCCCAAGAAGGCAACAGACCCTGAAATAGATGCCAATGTCTAGC CAAAAG	2176
	Qy	241	ACAAGAAGTGCTGGAAATTTTTGCCAAGCAGGAAAGAAGAACTCAGGTGAGCAGAAACACGCTT	300
	Db	2175	ACAAGAAGTGCTGGAAATTTTTGCCAAGCAGGAAAGAAGAACTCAGGTGAGCAGAAACACGCTT	2116
	Qy	301	TGCCTTTTC AATCAGCTTTTAACAGCGCTCTGAAGCTCCTTCCTATCATGCTAGTGCCTTCCTG	360
	Db	2115	TGCCTTTTC AATCAGCTTTTAACAGCGCTCTGAAGCTCCTTCCTATCATGCTAGTGCCTTCCTG	2056
	Qy	361	TTTTAGAGAGACTAACCAGAGACATTGAAGTCAGGTTAAAGCTGAATATATAACATTGCTGA	420
	Db	2055	TTTTAGAGAGACTAACCAGAGACATTGAAGTCAGGTTAAAGCTGAATATATAACATTGCTGA	1956
	Qy	421	AATGTTTTTCTTGTTGTTATTTTAACAGGGCTGAAGACATTATGAGAGAAAGACTGGAATAA	480
	Db	1995	AATGTTTTTCTTGTTGTTATTTTAACAGGGCTGAAGACATTATGAGAGAAAGACTGGAATAA	1936
	Qy	481	TCATAAGAAAGGAAAGAACTGTTCCAAGCTTGGGAAAAAGTGATTTTATCAGCAGTTAGT	540
	Db	1935	TCATAAGAAAGGAAAGAACTGTTCCAAGCTTGGGAAAAAGTGATTTTATCAGCAGTTAGT	1876
	Qy	541	GAGAGGAAGAAATATCAGAGAAAGTTTAGAGGAACACTTAAGACAAACCGAGTTAAGAGGG	600
	Db	1875	GAGAGGAAGAAATATCAGAGAAAGTTTAGAGGAACACTTAAGACAAACCGAGTTAAGAGGG	1816
	Qy	601	AAGGAAGAAAAATTAGGTTAAGAGGTTTACAAGAACAACTAGCCCCAGTCAGTGATGCCAG	660
	Db	1815	AAGGAAGAAAAATTAGGTTAAGAGGTTTACAAGAACAACTAGCCCCAGTCAGTGATGCCAG	1756
	Qy	661	CAGCGTGTTCTTCAGCGCCCTCTTACCGGGCAGGTGAAGACTTTAGAAAACAGTAGCAG	720
	Db	1755	CAGCGTGTTCTTCAGCGCCCTCTTACCGGGCAGGTGAAGACTTTAGAAAACAGTAGCAG	1696
	Qy	721	AGGAGATCTATGCATCCTATAGATTAAAGGAGCAAAGAAATCCTCTTAAATATTTCCA	780
	Db	1695	AGGAGATCTATGCATCCTATAGATTAAAGGAGCAAAGAAATCCTCTTAAATATTTCCA	1636
	Qy	781	TGAAGCTCTGSAATGCAAAACCGATGTGCTCTGTACTTTTAGCACATACCATTCATCTAC	840
	Db	1635	TGAAGCTCTGSAATGCAAAACCGATGTGCTCTGTACTTTTAGCACATACCATTCATCTAC	1576
	Qy	841	AGGTAGATTTC CCAACCAAATATATCCAGAGATGCCCTTTGTCAATCGGGTTATATACAGC	900
	Db	1575	AGGTAGATTTC CCAACCAAATATATCCAGAGATGCCCTTTGTCAATCGGGTTATATACAGC	1516
	Qy	901	CTTTGCGCTCTCTGAGTCAA TGATATTTACCATTTCCTCTGAGAAATCGAAAATCAATTTGG	960
	Db	1515	CTTTGCGCTCTCTGAGTCAA TGATATTTACCATTTCCTCTGAGAAATCGAAAATCAATTTGG	1456
	Qy	961	GGAGGGGACATTTAGAAAAAGAAATCAAAGTGCTCATGGATTAATCAAAATCTCTCAATAGTT	1020
	Db	1455	GGAGGGGACATTTAGAAAAAGAAATCAAAGTGCTCATGGATTAATCAAAATCTCTCAATAGTT	1396
	Qy	1021	GCAGTTATTCAGATGGCCAAAGGAAAAATAAAGTCATTAGATAGGGTTGGTAGAATTTAG	1080
	Db	1395	GCAGTTATTCAGATGGCCAAAGGAAAAATAAAGTCATTAGATAGGGTTGGTAGAATTTAG	1336
	Qy	1081	AACATGCTGTTTTTCAGGTTTATGGTCTTTTTTTTTTTTTTTTTTTTTTTTTTAAATAGGGAATG	1140
	Db	1335	AACATGCTGTTTTTCAGGTTTATGGTCTTTTTTTTTTTTTTTTTTTTTTTTTTAAATAGGGAATG	1276
	Qy	1141	TGTTTGGTGACAGCCAAATGTCATTC AAAAAGCTCTCTCTTTTCCCTGGTCAGTCATGTG	1200
	Db	1275	TGTTTGGTGACAGCCAAATGTCATTC AAAAAGCTCTCTCTTTTCCCTGGTCAGTCATGTG	1216
	Qy	1201	CTGGGACAGACAAGGGATCTGGATTAGGCAACATCATAGAGTGTCTCTGAGCTGCTCTTT	1260
	Db	1215	CTGGGACAGACAAGGGATCTGGATTAGGCAACATCATAGAGTGTCTCTGAGCTGCTCTTT	1156
	Qy	1261	GGTGATAACCCCTTCCAAATCTAAACTTTTTTGGAAATTC AAAGCTCAAGGCTCAAGGAGGAACCT	1320

Db	1155	GGTGATAACCCCTTCCAAATCCTAAACCTTTTGGAAATTCACARCTCAAGAGGAGAAACCT	1099
Qy	1321	ACTCTCTGATACACATGTTCTGCATTTTCTATCATGGTCTATGAAACATTTCTCTTA	1380
Db	1095	ACTCTCTGATACACATGTTCTGCATTTTCTATCATGGTCTATGAAACATTTCTCTTA	1036
Qy	1381	GAAATCCAGTGGCAAGAGTCTATGATTAAGTGTCTGAGCTCAGGCCAGGCAGTCA	1440
Db	1035	GAAATCCAGTGGCAAGAGTCTATGATTAAGTGTCTGAGCTCAGGCCAGGCAGTCA	976
Qy	1441	GAACTACTCTGAGTATTTACTACTCATTTGTGGGCAGCCCTCAGCTATCGGTTCTTTC	1500
Db	975	GAACTACTCTGAGTATTTACTACTCATTTGTGGGCAGCCCTCAGCTATCGGTTCTTTC	916
Qy	1501	ACACCTGCTATGAGATGATCCATATTTATGCTCGAGGCCAGCTATGCTCCCCACGAGA	1560
Db	915	ACACCTGCTATGAGATGATCCATATTTATGCTCGAGGCCAGCTATGCTCCCCACGAGA	856
Qy	1561	TCAGTCTTGAACCTAAACCTGGAATTTTATGAGTGTATTTATATGCAACTATTAATCA	1620
Db	855	TCAGTCTTGAACCTAAACCTGGAATTTTATGAGTGTATTTATATGCAACTATTAATCA	796
Qy	1621	ACATTACAGTCTCCCTCTGTATTTCTCTGTAAACATTTAGGCTGCGAAAAAARAA	1680
Db	795	ACATTACAGTCTCCCTCTGTATTTCTCTGTAAACATTTAGGCTGCGAAAAAARAA	736
Qy	1681	ATCTTTTTAAAAATAATGCCATAAAGTATTTGCTCTGGCCCTACTGTATGCTTCTTTC	1740
Db	735	ATCTTTTTAAAAATAATGCCATAAAGTATTTGCTCTGGCCCTACTGTATGCTTCTTTC	676
Qy	1741	TTTTTCTCTTTCACTAAGTCACCGTCAATTTATTAAGATGGCCATAACTATTTCAA	1800
Db	675	TTTTTCTCTTTCACTAAGTCACCGTCAATTTATTAAGATGGCCATAACTATTTCAA	616
Qy	1801	ACCTATGCTGAGTTCCCTCAAGCAGGGTCACATAGTATGAAGTGTGGGATGGGGTACG	1860
Db	615	ACCTATGCTGAGTTCCCTCAAGCAGGGTCACATAGTATGAAGTGTGGGATGGGGTACG	556
Qy	1861	GAGAACCCAGACAACCTCTAGTTTATTTAAACCTGTATTTACTGCCACTTTCCCTTTA	1920
Db	555	GAGAACCCAGACAACCTCTAGTTTATTTAAACCTGTATTTACTGCCACTTTCCCTTTA	496
Qy	1921	GACTTGACCATATGACCCCTCGTCCCAATTTCTAAGCATAGGGCAGGCCTTATTTTACA	1980
Db	495	GACTTGACCATATGACCCCTCGTCCCAATTTCTAAGCATAGGGCAGGCCTTATTTTACA	436
Qy	1981	ATGCTATATAGATATCACTCAGGTTTTATCAAGAGTTCGGCGCGTGTGGAAGTTCC	2040
Db	435	ATGCTATATAGATATCACTCAGGTTTTATCAAGAGTTCGGCGCGTGTGGAAGTTCC	376
Qy	2041	AAACAGATTTCAGTTTGTGTTGTCAGATTTCTAAATTTTACATGTTTCTTTTGCCAAAG	2100
Db	375	AAACAGATTTCAGTTTGTGTTGTCAGATTTCTAAATTTTACATGTTTCTTTTGCCAAAG	316
Qy	2101	GTGATTTTTTTAAATAACATTTGTTTCTCTATCTCTTTATAGTTCGAGACCAT	2160
Db	315	GTGATTTTTTTAAATAACATTTGTTTCTCTATCTCTTTATAGTTCGAGACCAT	256
Qy	2161	GAGAAACAGCGTCAAATCATCTTTTTCATGATCCCAAGCTGAAAGGCAAGCCCTCCAG	2220
Db	255	GAGAAACAGCGTCAAATCATCTTTTTCATGATCCCAAGCTGAAAGGCAAGCCCTCCAG	196
Qy	2221	CGCTTATGACCCCAACACCGACCATTTGGTGACAGACCTTCGGGCGCTGTCTGAAGCC	2280
Db	195	CGCTTATGACCCCAACACCGACCATTTGGTGACAGACCTTCGGGCGCTGTCTGAAGCC	136
Qy	2281	ATAGCCTCCACGAGAGCC	2299
Db	135	ATAGCCTCCACGAGAGCC	117

LOCUS AY434104 10006 bp DNA linear PRI 20-OCT-2003
DEFINITION Homo sapiens endothelin 1 (EDN1) gene, complete cds.
ACCESSION AY434104
VERSION AY434104.1 GI:37654537
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (bases 1 to 10006)
Rieder, M.J., Livingston, R.J., Daniels, M.R., Chung, M.-W.,
Miyamoto, K.E., Nguyen, C.P., Nguyen, D.A., Poel, C.L., Robertson, P.D.,
Schackwitz, W.S., Sherwood, J.K., Wittrak, L.A. and Nickerson, D.A.
Direct Submission
TITLE Submitted (09-OCT-2003) Genome Sciences, University of Washington,
JOURNAL 1705 NE Pacific, Seattle, WA 98195, USA
COMMENT To cite this work please use: NIEHS-SNPs, Environmental Genome
Project, NIEHS ES15478, Department of Genome Sciences, Seattle, WA
(URL: <http://egp.gs.washington.edu>).
FEATURES
source
1. 10006
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
102..107
/frequency="0.99"
/replace=""
variation
318
/frequency="0.26"
/replace="g"
variation
323
/frequency="0.01"
/replace="t"
variation
389
/frequency="0.01"
/replace="t"
variation
479
/frequency="0.03"
/replace="a"
variation
481
/frequency="0.03"
/replace="a"
variation
507
/frequency="0.26"
/replace="g"
variation
596
/frequency="0.01"
/replace=""
variation
678
/frequency="0.01"
/replace="c"
variation
927
/frequency="0.18"
/replace="c"
gene 1947..8080
/gene="EDN1"
mRNA join(1947..2214,3862..4030,5464..5619,5784..5927,
7486..8080)
/gene="EDN1"
/product="endothelin 1"
variation 2020
/gene="EDN1"
/frequency="0.83"
/replace=""
CDS join(2151..2214,3862..4030,5464..5619,5784..5927,
7486..7591)
/gene="EDN1"
/codon_start=1
/product="endothelin 1"
/protein_id="AAQ96600.1"
/db_xref="GI:37654538"
/translation="WDYLLIMFSLIFVACQAPETAVLGAFLSAVGENGGEKTPSPPP
WLRKRKRCSSLSMDKECVYFCHLDIITWNTPEHVVPYGLGSPRSKRALENLLPYKA

TDRENRCOASQKKQKWNFCQAGKELRAEDIMEKWNHKKGKDCSLGKKCIYQOL
VRGKIRSRSSSEHLRQTRSETMRNSVKSSFHDPLKKGKPSRRVYVTHRAHW"
2222 /gene="EDN1"
/frequency="0.01"
/replace="t"
2400
/gene="EDN1"
/frequency="0.47"
/replace="t"
2465 /gene="EDN1"
/frequency="0.06"
/replace="g"
2476 /gene="EDN1"
/frequency="0.03"
/replace="a"
2523 /gene="EDN1"
/frequency="0.01"
/replace="c"
3110 /gene="EDN1"
/frequency="0.01"
/replace="a"
3270 /gene="EDN1"
/frequency="0.46"
/replace="a"
3274 /gene="EDN1"
/frequency="0.01"
/replace="t"
3336 /gene="EDN1"
/frequency="0.01"
/replace="g"
3631 /gene="EDN1"
/frequency="0.02"
/replace="a"
3683..3686 /gene="EDN1"
/frequency="0.93"
/replace=""
3788 /gene="EDN1"
/frequency="0.01"
/replace="c"
3813 /gene="EDN1"
/frequency="0.01"
/replace="c"
3816 /gene="EDN1"
/frequency="0.03"
/replace="a"
4060 /gene="EDN1"
/frequency="0.45"
/replace="t"
4277 /gene="EDN1"
/frequency="0.04"
/replace="a"
4449..4551 /rpt_family="MIR"
/rpt_type="dispersed"
4509 /gene="EDN1"
/frequency="0.25"
/replace="a"
variation
repeat_region
variation

```
variation 4821 /gene="EDN1"
/frequency="0.03"
/replace="g"
4925 /gene="EDN1"
/frequency="0.01"
/replace="t"
4956 /gene="EDN1"
/frequency="0.04"
/replace="t"
4992 /gene="EDN1"
/frequency="0.06"
/replace="t"
5076 /gene="EDN1"
/frequency="0.02"
/replace="t"
5271 /gene="EDN1"
/frequency="0.01"
/replace="t"
5425 /gene="EDN1"
/frequency="0.23"
/replace="c"
5445 .5446 /gene="EDN1"
/frequency="0.99"
/replace="n"
5548 /gene="EDN1"
/frequency="0.12"
/replace="a"
6094 .6443 /gene="EDN1"
/region not scanned for variation
6751 /gene="EDN1"
/frequency="0.08"
/replace="a"
6793 /gene="EDN1"
/frequency="0.26"
/replace="a"
6825 /gene="EDN1"
/frequency="0.06"
/replace="g"
7104 /gene="EDN1"
/frequency="0.20"

Query_Match 98.8%; Score 2273.2; DB 9; Length 10006;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2298; Conservative 0; Mismatches 3; Indels 3; Gaps 2;

Qy 1 CTGAATGATGCTCCCAAGTGCTATGTCATGAGTCCCTTGTGTCGCCAGTGGAATAGTG 60
Db 5335 CTGAATGATGCTCCCAAGTGCTATGTCATGAGTCCCTTGTGTCGCCAGTGGAATAGTG 5394
Qy 61 TGTCCATGTCATTTTAAAGACTATTAACTACATAATAGTTTCTT--TCTCTCTTT 118
Db 5395 TGTCCATGTCATTTTAAAGACTATTAACTACATAATAGTTTCTTCTCTCTCTT 5454
Qy 119 GGATAATAGGACGCTGTTCCGTATGGACTTGGAGCCCTAGGTCAGAGAGAGCCCTGGA 178
Db 5455 GGATAATAGGACGCTGTTCCGTATGGACTTGGAGCCCTAGGTCAGAGAGAGCCCTGGA 5514
Qy 179 GAATTTACTTCCCAAGCAACAGACCGGTGAATAGATGCCAATGTGCTAGCCAAA 238
```

```
Db 5515 GAATTTACTTCCCAAGCAACAGACCGTGAGATAGATGCCAATGTGCTAGCCAAA 5574
Qy 239 AGACAAGAAGTGCTGGAAATTTTCCCAAGCAGAGAAAGAACTCAGTGAGCAGAAACACC 298
Db 5575 AGACAAGAAGTGCTGGAAATTTTCCCAAGCAGAGAAAGAACTCAGTGAGCAGAAACACC 5634
Qy 299 TTTGCTTTTCAATCAGTTTAAACAGCCTCCTGAACTCCTCTATCATGTTGCTGCTTCC 358
Db 5635 TTTGCTTTTCAATCAGTTTAAACAGCCTCCTGAACTCCTCTATCATGTTGCTGCTTCC 5694
Qy 359 TGTTTAGAGAGACTAACAGAGACATTTGAAAGTCAGGTTAAAGCTGAATATATCATTTGCT 418
Db 5695 TGTTTAGAGAGACTAACAGAGACATTTGAAAGTCAGGTTAAAGCTGAATATATCATTTGCT 5754
Qy 419 GAAATGTTTTTCTTGTGTTATTTTAAACAGGCTTGAAGACATTTATGAGAAAGCTGGAAT 478
Db 5755 GAAATGTTTTTCTTGTGTTATTTTAAACAGGCTTGAAGACATTTATGAGAAAGCTGGAAT 5814
Qy 479 AATCATAGAAAGAAAGACTGTTTCAAGCTTTGGGAAAGTGTATTTATCATGAGCTTGA 538
Db 5815 AATCATAGAAAGAAAGACTGTTTCAAGCTTTGGGAAAGTGTATTTATCATGAGCTTGA 5874
Qy 539 GTGAGAGAAAGAAAGTTCAGAGAAAGTTCAGAGAAACACCTAAGACAAACCCAGGTAAGAG 598
Db 5875 GTGAGAGAAAGAAAGTTCAGAGAAAGTTCAGAGAAACACCTAAGACAAACCCAGGTAAGAG 5934
Qy 599 GGAAGGAAGAAAGTTCAGAGAAAGTTCAGAGAAACCAACTAGCAGTCAAGTGTATTTATCATGAGCTTGA 658
Db 5935 GGAAGGAAGAAAGTTCAGAGAAAGTTCAGAGAAACCAACTAGCAGTCAAGTGTATTTATCATGAGCTTGA 6594
Qy 659 AGCAGCCTGTTCTCCAGCCTTCTTACCGGCGAGGTGAAGACATTAGAAACAGTAC 718
Db 5995 AGCAGCCTGTTCTCCAGCCTTCTTACCGGCGAGGTGAAGACATTAGAAACAGTAC 6054
Qy 719 AGAGGAGATCTATGCATCTCTATAGATTTAAAGGAGCAAAAGAAATCCCTCTTAAATATTTTC 778
Db 6055 AGAGGAGATCTATGCATCTCTATAGATTTAAAGGAGCAAAAGAAATCCCTCTTAAATATTTTC 6114
Qy 779 CATGAAGCTCTGGAATGCAACCGATGCTCTCTGTACTTTTACGACATACCATTTTCACT 838
Db 6115 CATGAAGCTCTGGAATGCAACCGATGCTCTCTGTACTTTTACGACATACCATTTTCACT 6174
Qy 839 ACAGGTAGATTTCCCAACCAAAATATATCCAGAGATGCCCTTTGTCTATTGGTTATATACA 898
Db 6175 ACAGGTAGATTTCCCAACCAAAATATATCCAGAGATGCCCTTTGTCTATTGGTTATATACA 6234
Qy 899 GCCTTTGCCCTCTCTGAGTCAATGTTATACCACTTTCCCTGAGAAATCGAAATCAATTT 958
Db 6235 GCCTTTGCCCTCTCTGAGTCAATGTTATACCACTTTCCCTGAGAAATCGAAATCAATTT 6294
Qy 959 GGGGAGCGGACATTTAGAAAAAGAAATCAAAAGTGTATGATTAATCAAAATTTTCAATAG 1018
Db 6295 GGGGAGCGGACATTTAGAAAAAGAAATCAAAAGTGTATGATTAATCAAAATTTTCAATAG 6354
Qy 1019 TTGCATTTATTCAGATGGCCAAAGGAAATAAAGTCAATTAGATAGGTTGGTGAATTT 1078
Db 6355 TTGCATTTATTCAGATGGCCAAAGGAAATAAAGTCAATTAGATAGGTTGGTGAATTT 6414
Qy 1079 AGAACATGCTGTTTTTTCAGGTTTATGCTTTTTTTTTTTTTTTTTTTTTTAAATAGGAAA 1138
Db 6415 AGAACATGCTGTTTTTTCAGGTTTATGCTTTTTTTTTTTTTTTTTTTTTTAAATAGGAAA 6474
Qy 1139 TGTGTTTTGTGTCAGAGCCCAATCTCATTTCCAAAAGCTCTCTCTTTTCTGTGTCAGTCAATG 1198
Db 6475 TGTGTTTTGTGTCAGAGCCCAATCTCATTTCCAAAAGCTCTCTCTTTTCTGTGTCAGTCAATG 6534
Qy 1199 TGTGTCGAGAGAGAGGATCTGGAATAGGCAACATCATAGAGTTGCTCTGAGCTGCTCT 1258
Db 6535 TGTGTCGAGAGAGAGGATCTGGAATAGGCAACATCATAGAGTTGCTCTGAGCTGCTCT 6594
Qy 1259 TGTGTCGATTAACCTTCCCAATCTTAAATCTTTTGGAAATTTCAAAAGTCAAAAGGAGAAAC 1318
Db 6595 TGTGTCGATTAACCTTCCCAATCTTAAATCTTTTGGAAATTTCAAAAGTCAAAAGGAGAAAC 6654
```

QY	1319	CTACTCTCTGATCACCACATGTTCTGCAATTTTCTATCATGTCATGAACTCTCT	1378
Db	6655	CTACTCTCTGATCACCACATGTTCTGCAATTTTCTATCATGTCATGAACTCTCT	6714
QY	1379	TAGAAATCCAGTGGCAAGAGTTCTATGATTAAGTGTTCGAGCTCAGGCCAGGAGTC	1438
Db	6715	TAGAAATCCAGTGGCAAGAGTTCTATGATTAAGTGTTCGAGCTCAGGCCAGGAGTC	6774
QY	1439	ATGAACTACTCTGAGTGTATTTACTACTGATTTCTGCGGAGCTCAGCTATCGGTTTCT	1498
Db	6775	ATGAACTACTCTGAGTGTATTTACTACTGATTTCTGCGGAGCTCAGCTATCGGTTTCT	6834
QY	1499	TCACACCTGCTTATGAGAGTATCCATATTTATGTCGAGGCCAGTAACTCTCCCAAGA	1558
Db	6835	TCACACCTGCTTATGAGAGTATCCATATTTATGTCGAGGCCAGTAACTCTCCCAAGA	6894
QY	1559	GATCAGTTTCTGAACCTACCTGGAATTTTATGCGTTTATATGCCAATTAAT	1618
Db	6895	GATCAGTTTCTGAACCTACCTGGAATTTTATGCGTTTATATGCCAATTAAT	6954
QY	1619	CAACATTACAGTCTTCTCCTCTGATTTCTCTGTAAACATTAAGGCTGCAAAAAAAA	1678
Db	6955	CAACATTACAGTCTTCTCCTCTGATTTCTCTGTAAACATTAAGGCTGCAAAAAAAA	7014
QY	1679	AAATCTTTTAAATAATATGCAATAAAGTATTTGCTCTGGGCTACTGATGCTTTT	1738
Db	7015	AAATCTTTTAAATAATATGCAATAAAGTATTTGCTCTGGGCTACTGATGCTTTT	7074
QY	1739	TCTTTTCTCTCTTCAACTAAGTACCGTCAATTTATTAAGATGCCAATTAATCA	1798
Db	7075	TCTTTTCTCTCTTCAACTAAGTACCGTCAATTTATTAAGATGCCAATTAATCA	7134
QY	1799	AAACCTATGCTGAGTTCCTCAAGCAGGGTCACTAGTGTGAAGTGGGATGGGCTA	1858
Db	7135	AAACCTATGCTGAGTTCCTCAAGCAGGGTCACTAGTGTGAAGTGGGATGGGCTA	7194
QY	1859	CGAAGAAACCGAACAATCTAGTTTATTTAAACCTGTATTTACTGCCACCTCCCT	1918
Db	7195	CGAAGAAACCGAACAATCTAGTTTATTTAAACCTGTATTTACTGCCACCTCCCT	7254
QY	1919	TAGACTTGACCATATGACCCCTCGCTCCCATTTCTAAGCATAGGGCAGGCTTTT	1978
Db	7255	TAGACTTGACCATATGACCCCTCGCTCCCATTTCTAAGCATAGGGCAGGCTTTT	7314
QY	1979	CAATGGTATAGATATCATCTGAGTGTATTTATCAAGAGTTTGGGCGGTGGAAGTC	2038
Db	7315	CAATGGTATAGATATCATCTGAGTGTATTTATCAAGAGTTTGGGCGGTGGAAGTC	7374
QY	2039	ACAACCAAGTTCAGGTTTGTGTGCGCAGATTTCTAATTTTACATGTTTCTTTGCCAA	2098
Db	7375	ACAACCAAGTTCAGGTTTGTGTGCGCAGATTTCTAATTTTACATGTTTCTTTGCCAA	7434
QY	2099	GGGTGA- TTTTGTAAATCAATTTGTTTCTCTATCTGCTTTATAGTGGGAGAC	2157
Db	7435	GGGTGATTTTTTAAATAAATTTGTTTCTCTATCTGCTTTATAGTGGGAGAC	7494
QY	2158	CATGAGAAACCGGTCAATATCTTTTATGATCCCAAGCTGAAAGGCAAGCCCTCCAG	2217
Db	7495	CATGAGAAACCGGTCAATATCTTTTATGATCCCAAGCTGAAAGGCAAGCCCTCCAG	7554
QY	2218	AGAGCGTTATGACCCCAACCGAGCAATTTGGTGAAGACCTTTGGGGCTGCTGAA	2277
Db	7555	AGAGCGTTATGACCCCAACCGAGCAATTTGGTGAAGACCTTTGGGGCTGCTGAA	7614
QY	2278	GCCATAGCTCCACCGAGAGCCCT	2301
Db	7615	GCCATAGCTCCACCGAGAGCCCT	7638
RESULT 4			
HUMEDN1B		12461 bp	DNA
LOCUS			linear PRI 07-NOV-1994

DEFINITION Homo sapiens endothelin-1 (EDN1) gene, complete cds.

ACCESSION J05008

VERSION J05008.1 GI:340555

KEYWORDS endothelin-1; vasoconstrictor.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 12461)

AUTHORS Inoue,A., Yanagisawa,M., Takuwa,Y., Mitsui,Y., Kobayashi,M. and Masaki,T.

TITLE The human preproendothelin-1 gene. Complete nucleotide sequence and regulation of expression

JOURNAL J. Biol. Chem. 264 (25), 14954-14959 (1989)

MEDLINE 89359303

PUBMED 2670930

COMMENT Original source text: Homo sapiens (clone library: lambda-EMBL3)

DNA.

FEATURES

source

Location/Qualifiers

1..12461

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/map="6p24-p23"

/clone_lib="lambda-EMBL3"

98..383

/rpt_family="Alu"

339..745

/bound_moiety="acute phase reactant regulatory element"

979..1039

/note="2-DNA region; putative"

2183..2188

/bound_moiety="acute phase reactant regulatory element"

2951..2958

/bound_moiety="TPA/JUN"

3241..3248

/bound_moiety="TPA/JUN"

3316..3328

/bound_moiety="NF-1"

3499..3505

/bound_moiety="TPA/JUN"

3510..3515

/gene="EDN1"

/note="G00-119-861"

3577..3582

/gene="EDN1"

/note="G00-119-861"

3608..3639

/gene="EDN1"

/note="G00-119-861"

/number=1

join(3876..3939,5585..5753,7183..7338,7503..7646,

912..9317)

/gene="EDN1"

join(3876..3939,5585..5753,7183..7338,7503..7646,

912..9317,

/gene="EDN1"

/codon_start=1

/product="endothelin-1"

/protein_id="AAA52339.1"

/db_xref="GI:556203"

/db_xref="GDB:G00-119-861"

/translation="VDYLMIFFLLFVACQAPETAVLCAELSAVGENGKPTSPPP

WLRKRCSSLDKCEVYFCHLDIIMVTPHEVVPVGLGSPSSKRALENLLPTKA

TRENRCKCCKDKKCNFCQAGKELRAEDINEKDMNHNKGDCKLKKCIYQQL

VGRKTRRSSEELRQTRSETMRNSVKSFFHDPKDKGKPSRERYTHRAHW"

3876..3938

/gene="EDN1"

/note="G00-119-861"

join(5677..5753,7183..7219)

/gene="EDN1"

/product="endothelin-1"

/note="big form of endothelin-1; G00-119-861"

sig_peptide

mat_peptide

mat_peptide	5677..5739 /gene="EDN1" /product="endothelin-1" /note="small form of endothelin-1; G00-119-861"	Db	7296	ACAAGAAGTGTCTGGAATTTTTCACAGCAGGAAAGAACTCAGGTGAGCAGAAACACCTT	7355
intron	3940..5584 /gene="EDN1" /note="G00-119-861" /number=1 /bound moiety="acute phase reactant regulatory element"	Qy	301	TGCTTTTCAATCAGTTTAAACAGCTCTCTGAACCTCTCTATCATGTTACTGCTTCTCTG	360
protein_bind	4612..4618 /gene="EDN1" /note="G00-119-861"	Db	7356	TGCTTTTCAATCAGTTTAAACAGCTCTCTGAACCTCTCTATCATGTTACTGCTTCTCTG	7415
exon	5585..5753 /gene="EDN1" /note="G00-119-861"	Qy	361	TTTTAGAGAGACTAAACAGAGACATTTGAAAGTCAAGGTAAAGCTGAATATAACATTCCTGA	420
intron	5754..7182 /gene="EDN1" /note="G00-119-861"	Db	7416	TTTTAGAGAGACTAAACAGAGACATTTGAAAGTCAAGGTAAAGCTGAATATAACATTCCTGA	7475
exon	7183..7338 /gene="EDN1" /note="G00-119-861"	Qy	421	AAATGTTTTCTTGTGTATTTTAAACAGGCTTCCAAAGCTTGGGAAAAAGTGTATTTATCAGCAGTTAGT	480
intron	7339..7502 /gene="EDN1" /note="G00-119-861"	Db	7476	AAATGTTTTCTTGTGTATTTTAAACAGGCTTCCAAAGCTTGGGAAAAAGTGTATTTATCAGCAGTTAGT	7535
exon	7503..7646 /gene="EDN1" /note="G00-119-861"	Qy	481	TCATAAGAAAGGAAAGACGTTCCTCAAGCTTGGGAAAAAGTGTATTTATCAGCAGTTAGT	540
intron	7647..9211 /gene="EDN1" /note="G00-119-861"	Db	7536	TCATAAGAAAGGAAAGACGTTCCTCAAGCTTGGGAAAAAGTGTATTTATCAGCAGTTAGT	7595
exon	9212..10443 /gene="EDN1" /note="G00-119-861"	Qy	541	GAGAGGAAGAAATCAGAGAGATTTCAAGGACCACTTAAGCAACCAAGTGAAGAGGG	600
intron	10444..10764 /gene="EDN1" /note="G00-119-861"	Db	7596	GAGAGGAAGAAATCAGAGAGATTTCAAGGACCACTTAAGCAACCAAGTGAAGAGGG	7655
exon	10765..10964 /gene="EDN1" /note="G00-119-861"	Qy	601	AAGGAGAAATAATTTAGGTAAAGGTTTCAAGGACCACTTAAGCAACCAAGTGAAGAGGG	660
intron	10965..11064 /gene="EDN1" /note="G00-119-861"	Db	7656	AAGGAGAAATAATTTAGGTAAAGGTTTCAAGGACCACTTAAGCAACCAAGTGAAGAGGG	7715
exon	11065..11264 /gene="EDN1" /note="G00-119-861"	Qy	661	CAGCTGTCTCTCCAGCCCTTCTTACCGGGCAGGTGAAGACTTGAAGAAACAGTACAG	720
intron	11265..11464 /gene="EDN1" /note="G00-119-861"	Db	7716	CAGCTGTCTCTCCAGCCCTTCTTACCGGGCAGGTGAAGACTTGAAGAAACAGTACAG	7775
protein_bind	11465..11664 /gene="EDN1" /note="G00-119-861"	Qy	721	AGGAGATCTATGCATCCTATAGATTAAAGAGGACCAAGAAATCCCTCTTAAATATTTCCA	780
protein_bind	11665..11864 /gene="EDN1" /note="G00-119-861"	Db	7776	AGGAGATCTATGCATCCTATAGATTAAAGAGGACCAAGAAATCCCTCTTAAATATTTCCA	7835
exon	11865..12064 /gene="EDN1" /note="G00-119-861"	Qy	781	TGAAGCTCTGGAATGCAACCGGATGCTCTGTATTTAGCAATACCAATTTCACTAC	840
intron	12065..12264 /gene="EDN1" /note="G00-119-861"	Db	7836	TGAAGCTCTGGAATGCAACCGGATGCTCTGTATTTAGCAATACCAATTTCACTAC	8895
protein_bind	12265..12464 /gene="EDN1" /note="G00-119-861"	Qy	841	AGGTAGATTTCCCAACCAAAATATATCCAGAGATGCTTTGTTCATTTGGTGTATATACAG	900
protein_bind	12465..12664 /gene="EDN1" /note="G00-119-861"	Db	7896	AGGTAGATTTCCCAACCAAAATATATCCAGAGATGCTTTGTTCATTTGGTGTATATACAG	9955
exon	12665..12864 /gene="EDN1" /note="G00-119-861"	Qy	901	CTTTGCTCTCTGAGTCAATGTATTTACACCTTTCCCTGAGAAATCGAAAAATCATTTGG	960
intron	12865..13064 /gene="EDN1" /note="G00-119-861"	Db	7956	CTTTGCTCTCTGAGTCAATGTATTTACACCTTTCCCTGAGAAATCGAAAAATCATTTGG	8015
protein_bind	13065..13264 /gene="EDN1" /note="G00-119-861"	Qy	961	GGAGCGGACATTTAGAAAAAGATCAAGTGTATGATGATGATGATGATGATGATGATGATG	1020
protein_bind	13265..13464 /gene="EDN1" /note="G00-119-861"	Db	8016	GGAGCGGACATTTAGAAAAAGATCAAGTGTATGATGATGATGATGATGATGATGATGATG	8075
exon	13465..13664 /gene="EDN1" /note="G00-119-861"	Qy	1021	GCAGTTATTCAGATGGCCCAAGGAAATAAAGTCAATTAGATAGGCTTGGTGAATTTAG	1080
intron	13665..13864 /gene="EDN1" /note="G00-119-861"	Db	8076	GCAGTTATTCAGATGGCCCAAGGAAATAAAGTCAATTAGATAGGCTTGGTGAATTTAG	8135
exon	13865..14064 /gene="EDN1" /note="G00-119-861"	Qy	1081	AACATGCTGTTTTTCAGGTTTATGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1135
intron	14065..14264 /gene="EDN1" /note="G00-119-861"	Db	8136	AACATGCTGTTTTTCAGGTTTATGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	8195
exon	14265..14464 /gene="EDN1" /note="G00-119-861"	Qy	1136	AAATGTTTTGTTGGTGAGAGCCATGTCATTTCAAAAAGCTCTCTCTCTCTCTCTCTCTCTCT	1195
intron	14465..14664 /gene="EDN1" /note="G00-119-861"	Db	8196	AAATGTTTTGTTGGTGAGAGCCATGTCATTTCAAAAAGCTCTCTCTCTCTCTCTCTCTCTCT	8255
exon	14665..14864 /gene="EDN1" /note="G00-119-861"	Qy	1196	ATGTGCTGGGACAGAGAGGATCTGGATTAGGCAACATCATAGATGCTGCTGAGCTGC	1255
intron	14865..15064 /gene="EDN1" /note="G00-119-861"	Db	8256	ATGTGCTGGGACAGAGAGGATCTGGATTAGGCAACATCATAGATGCTGCTGAGCTGC	8315
exon	15065..15264 /gene="EDN1" /note="G00-119-861"	Qy	1256	TCCTTGGTATTAACCTTCCAAATCTTAACTTTTGGAAATTCACAAAGCTCAAAAGGAGA	1315
intron	15265..15464 /gene="EDN1" /note="G00-119-861"	Db	8316	TCCTTGGTATTAACCTTCCAAATCTTAACTTTTGGAAATTCACAAAGCTCAAAAGGAGA	8375
exon	15465..15664 /gene="EDN1" /note="G00-119-861"	Qy	1316	AACCTACTCTCTGATCTACAGATGCTTGTGCAATTTTCTATCATGCTCTATGGAATCTTC	1375
intron	15665..15864 /gene="EDN1" /note="G00-119-861"	Db	8376	AACCTACTCTCTGATCTACAGATGCTTGTGCAATTTTCTATCATGCTCTATGGAATCTTC	8435

ORIGIN

Query Match	97.4%; Score 2241.6; DB 9; Length 12461;
Best Local Similarity	99.3%; Pred. No. 0;
Matches 2294; Conservative	1; Mismatches 5; Indels 10; Gaps 4;
Qy	1 CTGAAATGATGCTCCCAAGTGTATGTGATGAGCTCCCTGTGTGCGCCAGTGGAAATAGGTG 60
Db	7056 CTGAAATGATGCTCCCAAGTGTATGTGATGAGCTCCCTGTGTGCGCCAGTGGAAATAGGTG 7115
Qy	61 TGTCCATGTCTATTTAAAGACTATTAAATACACTAATATAGTTTCTCTCTCTCTCTCTCTCT 120
Db	7116 TGTCCATGTCTATTTAAAGACTATTAAATACACTAATATAGTTTCTCTCTCTCTCTCTCTCT 7175
Qy	121 ATAATAGGACGCTGTTCCGATGACCTTGAAGCCCTAGTTCACAGAGACCTTGGAGA 180
Db	7176 ATAATAGGACGCTGTTCCGATGACCTTGAAGCCCTAGTTCACAGAGACCTTGGAGA 7235
Qy	181 ATTTACTTCCCAAGGACACAGACCTGAAATAGATGCAATGTCTAGCCCAAAAG 240
Db	7236 ATTTACTTCCCAAGGACACAGACCTGAAATAGATGCAATGTCTAGCCCAAAAG 7295
Qy	241 ACAAGAAGTGTGGAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAAAACACCTT 300

QY 1376 TCTTAGAATCCAGTGGCAAGAGTTCTATGATTAAGAGTGTCTGAGCTCAGGCCAGCA 1435
DB 8436 TCTTAGAATCCAGTGGCAAGAGTTCTATGATTAAGAGTGTCTGAGCTCAGGCCAGCA 8495
QY 1436 GTCATGAATCTACTCTGAGTGTATTTACTACTGATTTGTGGGCGAGCCTCAGCTATCGGTT 1495
DB 8496 GTCATGAATCTACTCTGAGTGTATTTACTACTGATTTGTGGGCGAGCCTCAGCTATCGGTT 8555
QY 1496 TCTTCACACCTGCTTATGAGAGTATCCATATTTATGTCGAGGCCAGTAATGCTCCCA 1555
DB 8556 TCTTCACACCTGCTTATGAGAGTATCCATATTTATGTCGAGGCCAGTAATGCTCCCA 8614
QY 1556 CGAGATCAGTTCTTGAACCTACCTGGAATTTTATGGTGTATTTATGTCGAATATTA 1615
DB 8615 CGAGATCAGTTCTTGAACCTACCTGGAATTTTATGGTGTATTTATGTCGAATATTA 8674
QY 1616 AATCAACATACAGTCTTCCCTCTGTAATTTCTCTGTAAACATTAAGGCTGCAAAAA 1675
DB 8675 AATCAACATACAGTCTTCCCTCTGTAATTTCTCTGTAAACATTAAGGCTGCAAAAA 8734
QY 1676 AAAAAATCTTTTAAATAATTCGCAATAAGTATTTCTCTGGCCCTACTGTAATGCTC 1735
DB 8735 AAAAAATCTTTTAAATAATTCGCAATAAGTATTTCTCTGGCCCTACTGTAATGCTC 8794
QY 1736 TTTTCTTTTCTCTCTTTTCAACTAAGTCAACGTCACCTCAATTTTATTAAGATGCCAATACTAT 1795
DB 8795 TTTTCTTTTCTCTCTTTTCAACTAAGTCAACGTCACCTCAATTTTATTAAGATGCCAATACTAT 8854
QY 1796 TCAAAACCTATGCTCAGTCTCCTCAAGCAGGTCACATAGTGAAGAGTGGAGTGGG 1855
DB 8855 TCAAAACCTATGCTCAGTCTCCTCAAGCAGGTCGCTAGTGAAGAGTGGAGTGGG 8914
QY 1856 CTACCGAAGAAACCAAGAACTCTAGTGTATTTTAAACCTGTATTTACTGCCACCTTC 1915
DB 8915 CTACCGAAGAAACCAAGAACTCTAGTGTATTTTAAACCTGTATTTACTGCCACCTTC 8974
QY 1916 CTTAGACTTGACCATATGACCCCTCGCT - CCAATCTAAGCATAGGGGAGGCTTTAT 1974
DB 8975 CTTAGACTTGACCATATGACCCCTCGCT - CCAATCTAAGCATAGGGGAGGCTTTAT 9034
QY 1975 TTTACAAATGGTAATA - - - GATATCACTTGAAGTGTATTTATCAAGAGTGTGCGCGGGTGGT 2031
DB 9035 TTTACAAATGGTAATA - - - GATATCACTTGAAGTGTATTTATCAAGAGTGTGCGCGGGTGGT 9094
QY 2032 AAAGTTCAACACAGATTCAGGTTTGTGTCGAGATTCCTAATTTTACATGTTCTTT 2091
DB 9095 AAAGTTCAACACAGATTCAGGTTTGTGTCGAGATTCCTAATTTTACATGTTCTTT 9154
QY 2092 TGCCAAAGGTTGATTTTAAATAAATGATTTTCTCTTATCTTGTCTTTATAGGTC 2151
DB 9155 TGCCAAAGGTTGATTTTAAATAAATGATTTTCTCTTATCTTGTCTTTATAGGTC 9214
QY 2152 GGAGACCATGAGAAACAGCGTCAATATCTTTTATGATCCCAAGCTGAAGCAAGCC 2211
DB 9215 GGAGACCATGAGAAACAGCGTCAATATCTTTTATGATCCCAAGCTGAAGCAAGCC 9274
QY 2212 CTCGAGAGGCTTATGTCGCCCAACCCGAGCATTTGTTGACAGACCTTCGGGGCTG 2271
DB 9275 CTCGAGAGGCTTATGTCGCCCAACCCGAGCATTTGTTGACAGACCTTCGGGGCTG 9334
QY 2272 TCTGAAGCCATAGCCTCCACGGAGAGCCCT 2301
DB 9335 TCTGAAGCCATAGCCTCCACGGAGAGCCCT 9364

RESULT 5
A98526
LOCUS
DEFINITION
Sequence 1 from Patent WO9911778.
ACCESSION
A98526
VERSION
A98526.1
KEYWORDS
GI:6781612

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 12460)
AUTHORS Higenbottam, T. and McCormack, K.
TITLE ANTISENSE TREATMENT OF PULMONARY HYPERTENSION
JOURNAL Patent: WO 9911778-A 1 11-MAR-1999;
HIGENBOTTAM TIMOTHY (GB); MCCORMACK KEITH (GB)
FEATURES
source
1..12460
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 96.9%; Score 2230; DB 6; Length 12460;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2294; Conservative 0; Mismatches 5; Indels 11; Gaps 5;
QY 1 CTGAAATGATGCTCCCAAGTGTCTATGTGATGAGCTCTTGTGTGCCAGTGGAAATAGGTG 60
DB 7056 CTGAAATGATGCTCCCAAGTGTCTATGTGATGAGCTCTTGTGTGCCAGTGGAAATAGGTG 7115
QY 61 TGTCCATGTGTCTATTTAAAGACTATTAATTACACTAATATAGTTTCTTCTCTCTTTGG 120
DB 7116 TGTCCATGTGTCTATTTAAAGACTATTAATTACACTAATATAGTTTCTTCTCTCTTTGG 7175
QY 121 ATAAATAGGACCGTCTTCCGTATGACTTGGAAAGCCCTAGGTCCAAAGAGAGCCTTGGAGA 180
DB 7176 ATAAATAGGACCGTCTTCCGTATGACTTGGAAAGCCCTAGGTCCAAAGAGAGCCTTGGAGA 7235
QY 181 ATTTACTTCCCAAAAGGCAACAGACCGTGAATAGATGCGAATGTGCTAGCCAAAG 240
DB 7236 ATTTACTTCCCAAAAGGCAACAGACCGTGAATAGATGCGAATGTGCTAGCCAAAG 7295
QY 241 ACAAGAAGTGTGGAAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAAAACACCTT 300
DB 7296 ACAAGAAGTGTGGAAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAAAACACCTT 7355
QY 301 TGTCTTTCAATCAGTTTAAACGCTCTGAACTCCTTCTATCATGTGCTAGCTTCTCTG 360
DB 7356 TGTCTTTCAATCAGTTTAAACGCTCTGAACTCCTTCTATCATGTGCTAGCTTCTCTG 7415
QY 361 TTTTAGAGAGACTTAACAGAGACATTTGAAGTCAAGGTAAAGCTGAATATAATACATTGCTGA 420
DB 7416 TTTTAGAGAGACTTAACAGAGACATTTGAAGTCAAGGTAAAGCTGAATATAATACATTGCTGA 7475
QY 421 AATGTTTTTCTCTGTATTTTAAACAGGCTGAAGACATTTAGGAAAGACTGGAATAA 480
DB 7476 AATGTTTTTCTCTGTATTTTAAACAGGCTGAAGACATTTAGGAAAGACTGGAATAA 7535
QY 481 TCATAAGAGAAAGAAAGACTGTTCAGGTTTGGGAAAGTGTATTTATCAGCAGTTAGT 540
DB 7536 TCATAAGAGAAAGAAAGACTGTTCAGGTTTGGGAAAGTGTATTTATCAGCAGTTAGT 7595
QY 541 GAGAGGAAAGAAAGAAAGTTCAGAGGAAACCTTAAGACAAACCCAGGTAGAGGG 600
DB 7596 GAGAGGAAAGAAAGAAAGTTCAGAGGAAACCTTAAGACAAACCCAGGTAGAGGG 7655
QY 601 AAGGAAGAAAGAAATTAGTGAAGGTTTCAAGAGCAACTAGCCCCAGTCAGTGATGCCAG 660
DB 7656 AAGGAAGAAAGAAATTAGTGAAGGTTTCAAGAGCAACTAGCCCCAGTCAGTGATGCCAG 7715
QY 661 CAGCCTGTTCTCCAGCCCTTCTTACCCGGGCGAGTGAAGACTTTAGAAAACAGTAGCAG 720
DB 7716 CAGCCTGTTCTCCAGCCCTTCTTACCCGGGCGAGTGAAGACTTTAGAAAACAGTAGCAG 7775
QY 721 AGGAGATCTATGCTATGATTAAGAGGAAAGAAAGATCCCTCTTAATATTTCGA 780
DB 7776 AGGAGATCTATGCTATGATTAAGAGGAAAGAAAGATCCCTCTTAATATTTCGA 7835
QY 781 TGAAGCTCTGGAATGCAAAACCCAGTCTCTCTCTTCTAGCACATACCATTTCTACTAC 840

Db 7836 TGAAGCTCTGGAAATGCAAAACGATGCTCTCTGTACTTTTAGCATACATTTTCACTAC 7895
Qy 841 AGGTAGATTTCACACCAAAATATATCCAGAGATGCCCTTTGTCTATGTTGGTATATACAGC 900
Db 7896 AGGTAGATTTCACACCAAAATATATCCAGAGATGCCCTTTGTCTATGTTGGTATATACAGC 7955
Qy 901 CTTTGGCTCTGAGTCAATGATTTTACACCTTTCCCTGAGAAATCGAAATCATTTTGG 960
Db 7956 CTTTGGCTCTGAGTCAATGATTTTACACCTTTCCCTGAGAAATCGAAATCATTTTGG 8015
Qy 961 GGAGCGGACATTTAGAAAAGAAATCAAGTGTCTATGAGTAATCAAAATCTTTCAATAAGTT 1020
Db 8016 GGAGCGGACATTTAGAAAAGAAATCAAGTGTCTATGAGTAATCAAAATCTTTCAATAAGTT 8075
Qy 1021 GCAGTATTTCAGATGCCCAAGGAAATTAAGTCAATGATAGTGGTGGTAGAATTTAG 1080
Db 8076 GCAGTATTTCAGATGCCCAAGGAAATTAAGTCAATGATAGTGGTGGTAGAATTTAG 8135
Qy 1081 AACATGCTGTTTTTCAGGTTTATGCTC-----TTTTTTTTTTTTTTTTTTTTTAAATAGG 1135
Db 8136 AACATGCTGTTTTTCAGGTTTATGCTC-----TTTTTTTTTTTTTTTTTTTTTAAATAGG 8195
Qy 1136 AAATGTGTTTGTGAGAGCCCAATGTCATTCAAAAAGCTCTCTTTTCTGTCGTGTCAGTC 1195
Db 8196 AAATGTGTTTGTGAGAGCCCAATGTCATTCAAAAAGCTCTCTTTTCTGTCGTGTCAGTC 8255
Qy 1196 ATGTGCTGGGACAGAGAGGATCTGGATTAGGCAACATCATAGAGTTGCTCTGAGTGC 1255
Db 8256 ATGTGCTGGGACAGAGAGGATCTGGATTAGGCAACATCATAGAGTTGCTCTGAGTGC 8315
Qy 1256 TCTTTGGTGATAACCTTCCAAATCCTAAACTTTTGGAAATCAAACTCAAAAGGAGGA 1315
Db 8316 TCTTTGGTGATAACCTTCCAAATCCTAAACTTTTGGAAATCAAACTCAAAAGGAGGA 8375
Qy 1316 AACCTACTCTGATCTACCATGCTCTGATTTTCTGATTTTCTATCATGCTGTATGGAACCTC 1375
Db 8376 AACCTACTCTGATCTACCATGCTCTGATTTTCTGATTTTCTATCATGCTGTATGGAACCTC 8435
Qy 1376 TCTTAGAAATCCAGTGGCAAGAGTTCTATGATTAAGTGTCTGAGTCTAGGCCAGGCA 1435
Db 8436 TCTTAGAAATCCAGTGGCAAGAGTTCTATGATTAAGTGTCTGAGTCTAGGCCAGGCA 8495
Qy 1436 GTCATGAATCTCTGAGTATTTTACTACTGATTTTGGGCGAGCTCAGCTATCGTT 1495
Db 8496 GTCATGAATCTCTGAGTATTTTACTACTGATTTTGGGCGAGCTCAGCTATCGTT 8555
Qy 1496 TCTTCCACCTCTTATGAGAGTATCCATATTTATGCTCGAGGCGAGTATGCTCCCA 1555
Db 8556 TCTTCCACCTCTTATGAGAGTATCCATATTTATGCTCGAGG- CAGTATGCTCCCA 8614
Qy 1556 CGAGATCAGTTTCTGAATCACTGGAATTTTATAGGTTTATATGCTCAACTATTA 1615
Db 8615 CGAGATCAGTTTCTGAATCACTGGAATTTTATAGGTTTATATGCTCAACTATTA 8674
Qy 1616 AATCAACATTACAGTTCTTCCCTCTGATTTTCTCTGTAATAACATTAGGCTCGCAAAAA 1675
Db 8675 AATCAACATTACAGTTCTTCCCTCTGATTTTCTCTGTAATAACATTAGGCTCGCAAAAA 8734
Qy 1676 AAAAAATCTTTTAAAAATTAATGCCATAAAGTATTTCTCTGCGGCTCTGATGCTTC 1735
Db 8735 AAAAAATCTTTTAAAAATTAATGCCATAAAGTATTTCTCTGCGGCTCTGATGCTTC 8794
Qy 1736 TTTTCTTTTCTCTCTTTTCACTAAGTCAAGTCAATTTTAAAGTGGCCATTAATCTAT 1795
Db 8795 TTTT- TTTTCTCTCTTTTCACTAAGTCAAGTCAATTTTAAAGTGGCCATTAATCTAT 8853
Qy 1796 TCAAAACCTATGCTGAGTTCTTCAAGGAGGCTCAGATAGTGAAGTTGGATGGG 1855
Db 8854 TCAAAACCTATGCTGAGTTCTTCAAGGAGGCTCAGATAGTGAAGTTGGATGGG 8913
Qy 1856 CTACGGAGAAACCAAGCAACTCTAGTTTATTTAAACCTGTATTTACTGCCCACTCC 1915

8914 CTACGGAGAAACCAAGCAACTCTAGTTTATTTAAACCTGTATTTACTGCCACTTC 8973
Qy 1916 CTTTAGACTTGACCATATGACCCCTCGCT-CCCATTTCTAAGCATAGGGCAGCTTTATT 1974
Db 8974 CTTTAGACTTGACCATATGACCCCTCGCTCCCATTTCTAAGCATAGGGCAGCTTTATT 9033
Qy 1975 TTTACAAATGTAATA--GATATCACTTGAGGTTTATCAAAAGTTTGGCGGGTGGTG 2031
Db 9034 TTTACAAATGTAATAGATGATCACTTGAGGTTTATCAAAAGTTTGGCGGGTGGTG 9093
Qy 2032 AAAGTTTCAACACAGATTCAGGTTTGTGTCAGATTTCTAATTTTACATGTTCTTT 2091
Db 9094 AAAGTTTCAACACAGATTCAGGTTTGTGTCAGATTTCTAATTTTACATGTTCTTT 9153
Qy 2092 TGCCAAAGGTTGATTTTAAATTAACATTTGTTTCTTCTTATCTGTTTATTAGGTC 2151
Db 9154 TGCCAAAGGTTGATTTTAAATTAACATTTGTTTCTTCTTATCTGTTTATTAGGTC 9213
Qy 2152 GGAGACCATGAGAAACAGGTCCTCAATCATCTTTTTCATGATCCCAAGCTGAAAGGCAAGCC 2211
Db 9214 GGAGACCATGAGAAACAGGTCCTCAATCATCTTTTTCATGATCCCAAGCTGAAAGGCAAGCC 9273
Qy 2212 CTCAGAGAGCTTATGTGACCCACACCGAGACATTTGGTGACAGACCTTCGGGGCCTG 2271
Db 9274 CTCAGAGAGCTTATGTGACCCACACCGAGACATTTGGTGACAGACCTTCGGGGCCTG 9333
Qy 2272 TCTGAAGCCATAGCTCCACGAGAGCCCT 2301
Db 9334 TCTGAAGCCATAGCTCCACGAGAGCCCT 9363

RESULT 6
BD080748 12460 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION Antisense remedy of pulmonary hypertension.
ACCESSION BD080748
VERSION BD080748.1 GI:22626351
KEYWORDS JP 2001515011-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 12460)
AUTHORS Higebotom,T., McCormack,K. and Smith,A.
TITLE Antisense remedy of pulmonary hypertension
JOURNAL Patent: JP 2001515011-A 1 18-SEP-2001;
UNIVERSITY OF SHEFFIELD
COMMENT OS Homo sapiens (human)
PN JP 2001515011-A/1
PD 18-SEP-2001
PF 02-SEP-1998 JP 2000508789
PR 02-SEP-1997 GB 9718487.3
PI TIMOTHY HIGEBOTOM,KEITH MCCORMACK,ADRIAN SMITH PC
A61K31/708,A61M11/00,A61M15/00,A61P3/06,C12N15/09,C12N15/00 CC
Strandedness: Single;
CC Topology: Linear;
CC Antisense remedy of pulmonary hypertension
FH Key Location/Qualifiers
FT source 1..12460
FT /Organism='Homo sapiens (human)'.
FEATURES
source
1..12460
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

Query Match 96.9%; Score 2230; DB 6; Length 12460;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2294; Conservative 0; Mismatches 5; Indels 11; Gaps 5;
Qy 1 CTGAATGATGCTCCCAAGTCTATGTGATGAGCTCTTGTGCGCCACTGGAATAGGTC 60

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) 22388257 12477932 2 (bases 1 to 1425) Strausberg, R. Direct Submission Submitted (23-AUG-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgapbs@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc_mgc@hgrl.nih.gov Akter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dierich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
REMARK COMMENT	Clon distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: iRAK Plate: 78 Row: M Column: 2. Location/Qualifiers 1. 1425 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5405159" /tissue_type="Liver, adenocarcinoma" /clone_lib="NIH MGC_90" /lab_host="DH10B" /note="Vector: pCMV-SPORT6"
FEATURES source	Query Match Best Local Similarity 17.9%; Score 411.4; DB 9; Length 1425; Matches 423; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
ORIGIN	1878 TCTAGTTTATTAACCTGATTACTGCTCCCACTCCCTTACCATGACCATATGACC 1937 1313 TTTAGTTTATTAACCTGATTACTGCTCCCACTCCCTTACCATGACCATATGACC 1254 1938 CTGCTCCCATTTAAGCATAGGGGCGGCTTTATTTTACATGTTATAGATATCAC 1997 1253 CTGCTCCCATTTAAGCATAGGGGCGGCTTTATTTTACATGTTATAGATATCAC 1194 1998 TTGAGTTTATTAACAGAGTTGCGGCGGCTTGTGAAAGTTTACAA-CCAGATTGAGTTT 2056 1193 TTGAGTTTATTAACAGAGTTGCGGCGGCTTGTGAAAGTTTACAA-CCAGATTGAGTTT 1134 2057 TGTGTGCGAGATCTAATTTTACATGTTTCTTTTCCGAAAGGTTGATTTTAAAT 2116 1133 TGTGTGCGAGATCTAATTTTACATGTTTCTTTTCCGAAAGGTTGATTTTAAAT 1074 2117 AACATTGTTTCTTATCTTCTTATCTGTTTATTTAGGTCGGAGACCATGAGAAACGGTCAAA 2176 1073 AACATTGTTTCTTATCTTCTTATCTGTTTATTTAGGTCGGAGACCATGAGAAACGGTCAAA 1014 2177 TCATCTTTTTCATGATCCCAAGCTGAAGGCAAGCCCTCCAGAGAGCGTTATGTGACCCAC 2236 1013 TCATCTTTTTCATGATCCCAAGCTGAAGGCAAGCCCTCCAGAGAGCGTTATGTGACCCAC 954
QY	2237 AACCGAGCATTTGGTGACAGACCTTCGGGGCTGTCTGAAGCCATAGCTCCACGGAGA 2296 Db 953 AACCGAGCATTTGGTGACAGACCTTCGGGGCTGTCTGAAGCCATAGCTCCACGGAGA 894
QY	2297 GCCCT 2301 Db 893 GCCCT 889
RESULT 9	AC095692 259474 bp DNA linear HTG 13-NOV-2002 LOCUS Rattus norvegicus clone CH230-9D3, WORKING DRAFT SEQUENCE, 4 DEFINITION unordered pieces. AC095692 AC095692.5 GI:24940989 HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP. VERSION Rattus norvegicus (Norway rat) KEYWORDS SOURCE Rattus norvegicus ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. 1 (bases 1 to 259474) Muzny, D.M., Maric, M., Metker, M., Lee, A., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Devila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huiyik, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Liu, J., Kowals, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensheva, L., Louised, H., Lozano, R.J., Lui, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokeme, O., Okwuonu, G., Olarnpusagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfamkoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Shen, H., Sander, W., Savory, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Snead, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villanana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A. Direct Submission

TITLE

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
2 (bases 1 to 259474)
Worley, K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

3 (bases 1 to 259474)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On Nov 13, 2002 this sequence version replaced gi:2272932.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GDAS
Center clone name: CH230-9D3
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 226564 bases at least Q40
Consensus quality: 228265 bases at least Q30
Consensus quality: 229686 bases at least Q20
Estimated insert size: 232505; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 226873: contig of 226873 bp in length
* 226874 226973: gap of unknown length
* 226974 240221: contig of 13248 bp in length
* 240222 240321: gap of unknown length
* 240322 245800: contig of 5479 bp in length
* 245801 245900: gap of unknown length
* 245901 259474: contig of 13574 bp in length.

FEATURES
source
1..259474
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-9D3"
misc_feature
1..1313
/note="wgs_contig"

ORIGIN
Query Match 11.4%; Score 263.2; DB 2; Length 259474;
Best Local Similarity 53.4%; Pred. No. 3e-49;
Matches 1237; Conservative 0; Mismatches 813; Indels 266; Gaps 25;

Qy 13 TCCCAAGTGTATGTGATGAGCTCCTTGTGTGCCAGTGGATAGGTGTGTCATGTGTC 72
Db 82287 TCCGTACTTTGGTATCATGAGCTGTCCAAATGAATGCTCTCTGAATGTAGGCCCAAGGA 82228
Qy 73 ATTTTAAAGACTATTAAATACATAATATAGTTTCTTTCTCTCTTTGGGATAATAGCACG 132
Db 82227 ATTTGAAGATTATTAGTAGGCTAATATCTTTCTCTCTCTTTCCATTAACAGCGCG 82168
Qy 133 TTGTTCCGTATGAGCTTTGGAAGCC---CTAGGTCGAAGAGAGCTTGGAGAATTTACTTC 189
Db 82167 TCGTCCCGTATGAGCTAGGAAGCCCTTAGGCTAAGCGATCTTTGAAAGACTTACTTC 82108
Qy 190 CCACAAAGGCAACAGACCGTGAAAATAGATGCCAATGTCTGACGCAAAAAGACAGAAGT 249
Db 82107 CCACAAAGACACAGACCAAGGAAACAGATGCCAGTGTCTCACAAAAGACAGAAGT 82048
Qy 250 GCTGGAATTTTGGCAAGCAGGAAAAGAACTCAGGTGAGCAGAAAACACCTTTGCTTTCA 309
Db 82047 GCTGGAATTTTGGCAAGCAGACAAAGAACTCCGGTGAGTGAAGCCCTTGG----- 81995
Qy 310 ATCAGTTTAACAGCTCTGAACTCCTCTCTATCATGTGACTGCTTCTGTTTATAGAGA 369
Db 81994 ---AGTTATCCAGCTGCTGGCTGCTCCCTCTCA---GATGCTTTTCACTTTAGAGA 81942
Qy 370 GACTTAACAGAGACATTTGAAAGTTCAGGTTAAAGCTGAATATAACATTCCTGAAATGTTTT 429
Db 81941 AAGGAGCAGAGACATGACAGGCAAGGTGAATCTCACACAGCCATCGCTAAGTGTGTTT 81882
Qy 430 CTTTGTGTATTTTAAACAGGCTGAAGACATTTATGGAGAAAGACTGGAATATCATAGAA 489
Db 81881 CCTGTCTGTCTTCAACAGAGCCCAAGATACCATGCGAAGAAAGGCTTAAAGACTTCAAGAA 81822
Qy 490 AGGAAAAGACTGTTCCCAAGCTTGGGAAAAGTGTATTATCAGCAGTGTAGTGAGAGAA 549
Db 81821 GGGAAAACCTGTCCCAAGCTGGGAAAGAGTGTATCTATCAGCAGCTGTGTGAGGGAAG 81762
Qy 550 AAAAATCAGAAGAAGTTTACAGGAACACCTTAAGACAAACACAGGTAAAGGGAAGAA 609
Db 81761 AAAAATTAAGA----- 81752
Qy 610 AAATAGGTAGAGTTTCAAGAAACAACTAGCCCGAGTCAGTGATGCCAGGACCTGTT 669
Db 81751 ----AGGTAGAGGCTCATGGGAGCTCTCGGCCCTCGATACAGTGTCTAATGGTCCCT 81697
Qy 670 CCTCCAGCCCTTCTTACCCGGCAGGTGAAAGACTTTAGAAAACAGTAGCAGAGAGATCT 729
Db 81696 AGTCTTCC-----CACTGGGAGAAAGAAAGCTTGAGTACAGTAACAGAG---TCG 81648
Qy 730 ATGATCTCTATAGATTAAAGAGGACAAAGAAATCCCTCTTAAATATTTCCATGAAGCTCT 789
Db 81647 TCATGGTCTGTGGACTCAGAGGAGCAAGGGCTCCTCAGATACT-----GGGGCCT 81597
Qy 790 GGAATGCAAAACCGATGCTCTGTACTTTTAGCACATACCATTTCTATCTACAGTAGATT 849
Db 81596 TGAGTGAGAGGCTGCGCCACTGT-GTCTTAGCACACATTTGTTCACTATGATGAGC 81538
Qy 850 TCCCAACCAAAATATATCCAGAGATGCTTTGTGATTTGGTTATATACAGCCTTTGCTC 909
Db 81537 TTCCACACTA-----AGACAGCTCTAGAGTTTGAAGCATGCCAGTCTTTTCTCC 81489
Qy 910 TCTGAGTCAATGTATTATACCACTTTCCCTCAGAAATCGRAAATCAATTTGGGGAGCGAC 969
Db 81488 TCT--GGTCTTTCAATTTCTGCTTGGCTGGAGAACTGAAAAAT-----C 81447
Qy 970 ATTTAGAAAAGAAATCAAAAGTGTATGATAAATCAAAATCTTCAATAAGTTGCAATTT 1029
Db 81446 ATGCCAAAGAAAAGACATTTAAAAATATTTATATATGTTACTCAAGAGTGTAGTGAAT 81387
Qy 1030 CAGATGCCAAAGAAAATTAAGTCAATAGTAGGTTGGTAGAATTTAGACATGCTG 1089
Db 81386 CAAATGCCAAAGG-GAAATAAGCCCAATAGTAAGGTTGGTAGATTTAGACATGTTG 81328
Qy 1090 TTTTTCAGTTTATGCTCTTTT-----TTTTTTTTTTTTTAAATAGGGAATGTGTTGGTG 1149

```
Db 81327 TTTTTCAGGTTTA-----TTCCGTTACTATTGTTGAATATGAAATTTGAAGGCC 81276
QY 1150 CAGAGCCATGTCATTCACAAAAGCTCTCTCTTTTCCTGGTCAGTCATGCTGGACAC 1209
Db 81275 AGGTGTCATCTGGCTTGC-----TTCTCAGGCCC 81246
QY 1210 AGAAGGGATCTGGATTAGCAACATCATAGAGTTGCTCTGAGCTGCTCTTTGGTGATAAC 1269
Db 81245 AGGAGAGGCTGATAGGCAATGCCACAGGCTGCAACAGCAGCTTTT--TGGTAT 81188
QY 1270 CTTTCAATCTTAATCTTTTGGAAATTCACAGCTCAAGAGGAGAACTACTCTCTGA 1329
Db 81187 CTTTCAAGGCTAATCTTTTGAAGGAGGACCTTTGTTGAGATCTGTCTATAGGC 81128
QY 1330 TCTACCAATGTTCTGCATTTTCTATCATGCTCTATGAAACTCTCTTGAATACTCAG 1389
Db 81127 TCACCTCTA-----TCTCTAGAGTGTCCCAAGTCTATCCAG 81091
QY 1390 TGGCAAGAGTCTATGATTAAGTGTCTGAGCTCAGCCAGCAGTCATGAATCTT 1449
Db 81090 GGTCAAGGCTTCTGGGACCAAGTGTCTGAGCCCAATGT--GGCGTGGGAGGAGCTG 81032
QY 1450 CTGAGTTATTTACTACTGATTTGCGGG--CAGCTCAGCTATCGTTTCTTTCACACCTGC 1508
Db 81031 CTGACTTGGTTACTCTCTCTGTGGGACAGCCCTCTTAGCTCTCTCTTCCGACCCAC 80972
QY 1509 TTATGAGAGTATCCATATTTATGTCGCGAGGCGAGTAATGCTCCCGCAGAGATCAGTTTC 1568
Db 80971 CACTG-----TGGGGTCCCAACCAGGTTCACAGTTCACAGCTCAGCTGAGGC 80926
QY 1569 TGAATCACTGGAATTTTATGGTTTATGATGCAACTATTAATCAATCAATACA 1628
Db 80925 TCAAT-----ACCTTTTGTACTCTGTCTTCCATTAATCAATCAAGAGTGTCTCCA 80874
QY 1629 GTTCTTCCCTCTGTATTTCTCTGTAAACATTAGGCTGCAAAAAAATAATCTTTT 1688
Db 80873 GTTCTTATGTCCTCTCTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 80814
QY 1689 AAAAATAATGCAATAAGTATTTGCTCTGGGCTTACTGTATGCTTCTTTCTTTCTC 1748
Db 80813 NNNNNNNNNNACA-----CACACACACACACACACTGACATTCAC 80772
QY 1749 TCTTTTCAACTAAGTCACCGTCAATTTATTAAGATGGCCATACTATTCAAAACCTATGC 1808
Db 80771 AACCTTTAACTGTCACTAGCACTCGTTAGTCTCCCTGTAGGCTCCCTAATTTCT 80712
QY 1809 TGAGTTCTCAAGGAGGCTCACAATAGTGAAGGTTGGAGTGGGCTACGGAAGAAC 1868
Db 80711 CTCCATCCACCTGTAGCTTCCATTTCTCTGTGTGTTAGAAAGTGCAGAGGCTAGAT 80652
QY 1869 CAGAACACTCTAGTTTATTAACCTGTATTTACTGCCACCTCCCTTAGACTTGC 1928
Db 80651 -----GACCTACTTAAGCAGAACATTTGGAACTTAATCTCCTGTAGACATCT 80603
QY 1929 CATATGACCCCTCGCTCCCATTTAAGCATA-----GGGGCAGGCTTTATTTTACAATGG 1984
Db 80602 CACAGGCTTCTCTCCACAGCACCACAAACAGCATAGACAGCAGGCTTATTCACACATAG 80543
QY 1985 TAATAGATATCACTTGAGGTTTATCAAGAGTGTGGCG-----GGTGGTGA 2034
Db 80542 TGTAGCGCTGCTTGAGGATTTATCAAGAGGACCAACATGCTCAAACTGGTGAAG 80483
QY 2035 GTTCAACACCAATTCAGGTTTGTGTCGAGATCTTAATTTATACATGTTCTTTTGC 2094
Db 80482 GGCCAGGAACCTAAGACTTAGTGGCTTGTGCGAGATCTAATTTAGACATTTCTTAGC 80423
QY 2095 CAAAGGGTGAATTTTAAA-----ATAACATTTGTTTCTTCTATCTGCT 2141
Db 80422 CAACAGGTGATTTGTTCCATCTCTGTGAGTGTGTTGAAATATTTCTCTCTCTGTT 80363
QY 2142 TTATTAGTCTGGAGACCATGAGAACACGCTCAATCATCTTTTTCATGATCCCAAGCTGA 2201
```

```
Db 80362 TTAATAGTTGGAGGCCATCAGCAACAGCATCAAGACCTCTCTTTGAGTTGCAAAATTGA 80303
QY 2202 AAGCGAAGCCCTCCACAGAGCGTTATGTGACCAACCGACACATTTGGTGACAGACT 2261
Db 80302 AAGCGAATCTACAGAGACCGAGAGTTGATACACACCGACACATTTGACTACAGGCC 80243
QY 2262 TCGGGGCTGTGTGAAGCCATAGCCCTCCACGGAGAG 2297
Db 80242 CCGTGTGTTTGAAGCCATGACTTACATAGAGCG 80207

RESULT 10
HUMETN4
LOCUS Human endothelin (ETN) gene, complete cds. linear PRI 08-NOV-1994
DEFINITION M25380 J04819
ACCESSION M25380.1 GI:182257
VERSION M25380.1
KEYWORDS endothelin.
SEGMENT 4 of 4
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1330)
AUTHORS Bloch,K.D., Friedrich,S.P., Lee,M.E., Eddy,R.L., Shows,T.B. and
Quettermoue,T.
TITLE Structural organization and chromosomal assignment of the gene
encoding endothelin.
JOURNAL J. Biol. Chem. 264 (18), 10851-10857 (1989)
MEDLINE 89278161
PubMed 2659594
COMMENT Original source text: Human leukocyte DNA and human umbilical vein
endothelial cell cDNA to mRNA.
Draft entry and computer readable copy of sequence [1] kindly
submitted by K.D.Bloch (06-AUG-89).
FEATURES
Location/Qualifiers
source
1..1330
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="6p24-p23"
join(M25377:1:519..682,M25378:1:1..282,M25379:1:1..664,
1..206)
gene
/gene="EDN1"
join(M25377:1:519..582,M25378:1:101..269,
M25379:1:101..256,M25379:1:421..564,101..206)
CDS
/gene="EDN1"
note="endothelin prepropeptide"
/codon_start=1
/protein_id="AAA52407.1"
/db_xref="GI:182259"
/translation="MDYLLMLFSLFVACGAPETAVLGAEISAVGNGEKPTSPSP
WPLRSKCSLSMDKSCVYFCHLDIIIVNTPEHVYPYGLGSPRSKALSNLPTKA
TDRENRCCASOKKCNWFCAGKELRAEDIMEDKNHNKRGKDCSKLGHKCIYQOL
VGRKIRSEEHQRSTMRNSVRKSFDPKLGKPSRERYVTHRAHW"
prim_transcript <1..1330
/note="ENT mRNA and introns"
intron
<1..100
/gene="EDN1"
/note="ENT intron D"
101..206
/exon
/gene="EDN1"
/note="endothelin prepropeptide; G00-119-861"
/number=5
ORIGIN About 1.3 kb after segment 3.
Query Match 10.5%; Score 241; DB 9; Length 1330;
Best Local Similarity 99.6%; Pred. No. 4,5e-44;
Matches 252; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 2049 TCAGGTTTGTGTGCCAGATTCATATTTTACATGTTTCTTTTCCAAAGGCTGATTT 2108
Db 1 TCAGGTTTGTGTGCCAGATTCATATTTTACATGTTTCTTTTCCAAAGGCTGATTT 60
```

```

QY 2109 TTTTAAATAACATTTGTTTCTTCTATCTTCTTTATTAGTTCGAGACCATGAGAAACA 2168
Db 61 TTTTAAATAACATTTGTTTCTTCTATCTTCTTTATTAGTTCGAGACCATGAGAAACA 120
QY 2169 GCGTAAATCATCTTTTATGATCCCAAGCTGAAGGCAAGCCCTCCAGAGAGCGTTATG 2228
Db 121 GCGTAAATCATCTTTTATGATCCCAAGCTGAAGGCAAGCCCTCCAGAGAGCGTTATG 180
QY 2229 TGACCCACAACCGAGACATTTGGTGACAGACCTTCGGGGCCCTGTCTGAAGCAATAGCCCTC 2288
Db 181 TGACCCACAACCGAGACATTTGGTGACAGA-CTTGGGGCCCTGTCTGAAGCAATAGCCCTC 239
QY 2289 CACGGAGAGCCCT 2301
Db 240 CACGGAGAGCCCT 252

RESULT 11
AC131424
LOCUS
DEFINITION
Rattus norvegicus clone CH230-248L7, WORKING DRAFT SEQUENCE, 5
unordered pieces.
AC131424
AC131424.4 GI:25138078
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 237384)
Muzny, D., Marie, Metzker, M., Lee, A., Brannon, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Ayodeji, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhal, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Centek, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleaveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escoto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, J., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Gant, R., Garcia, A., Garner, T., Garza, M.,
Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kratt, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshelew, L., Louise, H., Lozada, R.J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,
Mawney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Moragan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwankelimeh, O., Okwuonu, G., Olarnpusagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
Puar, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Roze, R., Ruiz, S.J.,
Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,

```

```

Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wlaczek, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhauser, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 237384)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (22-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 237384)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 20, 2002 this sequence version replaced gi:23194834.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GUVH
Center clone name: CH230-248L7
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 202760 bases at least Q40
Consensus quality: 204531 bases at least Q30
Consensus quality: 205595 bases at least Q20
Estimated insert size: 206846; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* *
* 1 190059: contig of 190059 bp in length
* 190060 130159: gap of unknown length
* 190160 191371: contig of 1212 bp in length
* 191372 191471: gap of unknown length
* 191472 192425: contig of 2774 bp in length
* 192426 194345: gap of unknown length
* 194346 206754: contig of 12409 bp in length
* 206755 206854: gap of unknown length
* 206855 237384: contig of 30530 bp in length.
* Location/Qualifiers
* 1 .237384
* /organism="Rattus norvegicus"
FEATURES
source

```



```
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-248L7"
misc_feature
1..1192
  /note="wgs_end_extension
  clone_end:77"
misc_feature
3660..4806
  /note="wgs_end_extension
  clone_end:77"
misc_feature
5601..6569
  /note="clone_boundary
  clone_end:17
  site:
  end_sequence:B2193636"

ORIGIN
Query Match
Best Local Similarity 10.4%; Score 239; DB 2; Length 237384;
Matches 1228; Conservative 53.3%; Pred. No. 9.5e-44; Indels 260; Gaps 24;
13 TCCCAAGTGCTATGATCAGCTCCTGTGTGCCAGTGGAAATAGGTGTGTCATGTGC 72
42090 TCCGTACTTGGTATATGATGCTGCCAATGAATGCTCTCTGAATGTAGGCCCAAGGA 42149
73 ATTTTAAAGACTTAAATACACTAATATAGTTCTTCTCTCTTGGATATAGGACG 132
42150 ATTTGAAAGATTAATAGTAGGCTAATATCTTTCTCTCTCTTCCATACAGGCGG 42209
133 TTGTTCCGTATGCACTTGGAGCC---CTAGGTCCAGAGAGCCTTGGAGAAATTTACTTC 189
42210 TCGTCCGTATGGAATAGGAGCCCTCTAGGTCTAAGCGATCCTTGAAGACTTACTTC 42269
190 CCACAAAGCAACAGACCGTGAATAATAGATGCCAATGTGTAGCCAAAAGACAAAGT 249
42270 CCACAAAGCAACAGACCGTGAATAATAGATGCCAATGTGTAGCCAAAAGACAAAGT 42329
250 GCTGGAATTTTGGCAAGCAGGAAAGAACTCAGGTGAGCAGAAACACCTTTGCTTCA 309
42330 GCTGGAATTTTGGCAAGCAGCAGAAAGAACTCAGGTGAGTGGAAACCCCTTG----- 42382
310 ATCAGTTTAAAGCTCTCTGAACTCCTTCTATCATGTGATCTGCTTCTCTTTTAGAGA 369
42383 --GAGTTATCCAGCTTGTGCGCTCTCCTCTCA-----GATGCTTTTCACTTTAGAGA 42436
370 GACTAACAGACATTTGAAGTTCAGGTAAGAGTGAATATACATATGCTGAATGTTTTT 429
42437 AAGGAGCAGACATTTGAAGTTCAGGTAAGAGTGAATATCAACA--GCCATCGTAAAGTGT 42495
430 CCTTGCTGATTTTAAACAGGCTGAAGACATTTATGGAGAAAGCTGGAAATATCAAGAA 489
42496 CCTGCTGCTTCTCAGAGCCCAAGTACCATGAGAAAGGCGTAAAGACTTCAAGAA 42555
490 AGGAAAGACTTCTCAGAGCTTGGAAAGAGTGAATTTATCAGCAGTTAGTGAGGAGAG 549
42556 GGGAAAGCCCTGCTCCAGCTGGGAAAGAGTGAATTTATCAGCAGCTGGTGAGGAGAG 42615
550 AAAAATCAGAAAGTTTCAGAGGACACCTTAAGACAAACCCAGGTAAGAGGGAAGAGAA 609
42616 AAAAATCAGAAAGTTTCAGAGGACACCTTAAGACAAACCCAGGTAAGAGGGAAGAGAA 42625
610 AAATTAGTGAAGTTTCAAGAAACAACTAGCCAGTCAAGTATGCCAGCAGCTGTT 669
42626 -----AGGTAAGAGGCTCATGGGGACTCTCTGCGCTGGATACCAGTGTAAATGGCTCCCT 42680
670 CTTCCAGCCTCTTCTTACCGGCGAGTGAAGACATTTAGAAACAGTAGCAGAGAGATCT 729
42681 AGTCCTTCC-----CACTGGGAGAAAGAGCTTGAGTACAGTAAACAGAG-----TCG 42729
730 ATGCATCTCTATAGATTAAAGAGCAAGAAATCCCTCTTTAAATATTTCCATGAAGCTCT 789
42730 TCAATGCTGTGGAATCAGAGGAGCAAGAGGCTCTCAGACTACT-----GGGCGCT 42780
790 GGAATGCAACCGATGCTCTGTACTTTTAGCACATACCATTTTCACTACAGGTAGATT 849
```

```
Db 42781 TGAGTCAGACGCGCTGCCACTGT--GTCTTAGCACACATTTGTTCACTATGATGGAGC 42839
Qy 850 TCCCAACCAAAATATATCCAGAGATGCTTTCTCATGGGTTATATATACAGCTTTGCCCTC 909
Db 42840 TTCCACACTA-----AGACAGCTCTAGAGTTTGGCATGCCAGCTCTTTCTCC 42888
Qy 910 TCTGAGTCAATGATTTTACCCTTTCCCTGAGAAATCGAAATCATTTTGGGGAGCGAC 969
Db 42889 TCT--GGTCTTTCATTTTCTGCTTGGTGGAGAACTGAAAT-----C 42930
Qy 970 ATTTAGAAAAGAAATCAAGTCTCATGATATCAAAATCTTCAATAAGTTGCACTTATT 1029
Db 42931 ATGCCAAAGAAAAGACATTTAAATATTTATATGTTACTCCAGAGTGTCTAGTATT 42990
Qy 1030 CAGATGCCCAAGGAAATAAAGTCAATAGATAGGTTGTAGAAATTTAGAACATGCTG 1089
Db 42991 CAAATGCCCAAGG--GAATAAAGCCATTAGATAAGTTGGTGAGATTTAGAACATGTTG 43049
Qy 1090 TTTTTCAGGTTTATGCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1149
Db 43050 TTTTTCAGGTTT-----TTCGTTACTATTTGTTGAAATATGAAATTTGAAGGCCC 43101
Qy 1150 CAGAGCAATGTCTATTCACAAAAGCTCTCTCTTTCTGGTCAATGCTGTGGGACAG 1209
Db 43102 AGGTGTCACTTGGCTTGC-----TTCTGAGGCCC 43131
Qy 1210 AGAAGGATCTCGATAGGCAACATAGAGTGTCTGTAGCTGCTCTTTGGTGATTAAC 1269
Db 43132 AGGAGAGGCTGATAGGCAATGCCAGGGTGGCAACAGCAGCTTTT--TGGTGAT 43189
Qy 1270 CCTTCCAAATCTTAACTTTTGGAAATTCACAAGCTCAAAGGAGAACTACTCTCTGA 1329
Db 43190 CCTTCCAAAGGCTTAACTTTTGAACGTGAAGAGACCTTGTTCAGATCTCTATAGC 43249
Qy 1330 TCTACCACATGTTCTGCAATTTTCTATCATGTCTATGAAACTTCTCTTAGAAATCCAG 1389
Db 43250 TCACTCTA-----TCTCTAGAGTGTCCCAAGTCACTCCAG 43286
Qy 1390 TGGCAGAAATCTATGATTAAGTGTCTGAGCTCAGGCCAGGAGCATGAACTACTT 1449
Db 43287 GGTCAAGGCTTCTGGGACCAAGTGTCTGAGCCCAATGT--GGCCGTGGGAGGAGATG 43345
Qy 1450 CTGAGTTATTTACTACTGATTTTGTGGG--CAGCCTCAGCTATCGGTTTCTTCAACCTGC 1508
Db 43346 CTGACTTGGTTACTGCTCCTGTGTGGGACAGCCCTCTTAGCTCTCTCTCGCAC-- 43402
Qy 1509 TTATGAGATATCCATATTTATGCTGCGAGGCAATGCTCCCAAGAGATCAATTTTC 1568
Db 43403 -----CACCACTGTGGGTCCTCCACAGGTCAAGTTGC 43435
Qy 1569 TGAACCTAACCTGGAATTTTATGGGTTT--TTATTATGCCAACTATTAAATCAACATTAC 1627
Db 43436 ACAGCTGACCTGAGGCTCAATACCTTTTGTGTACTCTGTCTTCCATTAACTACGAAGTG 43495
Qy 1628 AGTTCTTCCCTCTGATTTTCTCTGTAAACATTTAGGCTGCAAAAAAATAAATCTTTT 1687
Db 43496 ATGCTCCAGTTCTTTATGTGTCTCTCATATGTGTGCGGTGCACAGACACACACA 43555
Qy 1688 TAAATAATTTGCATAAAGTATTTGCTCTGGGCTACTGTATGCTCTTTTCTTTTCT 1747
Db 43556 CACACACACACACACACTGACATGACACACCTTTTAACTGTCACTAGCACTCGCT 43615
Qy 1748 CTCTTTTCAACTAAGTCAACGCTCAATTTATTAAGATGGCCATAACTATTTCAAAACCTATG 1807
Db 43616 TAGGTCTCCCTGTAGGCTCCCTAAATTTTCTCTCCACTCCACCTGTTAGCTTCAATT----- 43670
Qy 1808 CTGAGTTCTCAGGAGGCTCAGATAGTATGAAAGTTGGATGGGTGGGTACGGAAGAAA 1867
Db 43671 -----TCTTCTGTGTTTGAAGAGTCCAGAGGTAGATGACGCTACTTTAAGCAG 43721
Qy 1868 CCAGAAACAATCTAGTTTATTTTAAACCTGTATTTTACTGCCCACCTTCCCTTAGACTTGA 1927
```



```
Db 43722 AACATTGGAACTTAACCTCCCTGTAGACATCTCACAGCTTCCTCTCCACAGACACAAA 43781
Qy 1928 CCATATGACCCCTCGCTCCCTCCTTAAGCATAGGCGCAGGCTTTATTTTACAAATGGTAA 1987
Db 43782 CAGCATAGACAGCAGCGGCTTATTCACACAATAGTCGTAGCCGCTGCTTTGAGGATTTATCA 43841
Qy 1988 TAGATATCACTTTGAGGTTTATCAAAAGAGTTCCGCGGGTGTGAAAGTTTACACACCAGA 2047
Db 43842 AAGAGCCACCAATGCTACAC-----TGGTGCAAGGGCCAGGAACTAA 43887
Qy 2048 TTCAGTTTTGTTGTCAGATCTTAATTTTACATGTTTCCTTTGCCAAAGGGTGATTT 2107
Db 43888 GACTTAGTCGGTTGTCGAGATCTTAATTTTACATGTTTCCTTTGCCAAAGGGTGATTT 43947
Qy 2108 TTTTAAATAACATTG-----TTTTCTCTATCTCTCTTTATAGTCGGA 2154
Db 43948 GTTTCATCTCTGAGTGTTGTTGAATATTTCTCTCTCTCTTTATAGTTGGA 44007
Qy 2155 GACCATGAGAAACAGCGTCAAAATCATCTTTTCATGATCCAGCTGAAAGCGAACCCCTC 2214
Db 44008 GGCATCAGCAACAGCATCAAGACCTCCTTTTCGAGTTGCAAAAGTTGAAAGCGGAATCTA 44067
Qy 2215 CAGAGCGTTATGTAGCCCAACAGCAGCATTGGTGACAGACTTCGGGGCCTGTCT 2274
Db 44068 CAGAGCCAGAAGTTGATACAAACCGAGCATTGACTACAGAGCCCGGTGTTTG 44127
Qy 2275 GAAGCATAGCTCCACGGAGAG 2297
Db 44128 GAAGCATGACTTACATAGAGCG 44150

RESULT 12
AF130760 266 bp DNA linear MAM 03-FEB-2000
LOCUS
DEFINITION
EQUUS caballus chromosome 20 EDN1 gene, partial sequence.
ACCESSION
AF130760
VERSION
AF130760.1 GI:6855655
KEYWORDS
EQUUS caballus (horse)
ORGANISM
EQUUS caballus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
REFERENCE
1 (bases 1 to 266)
CAetano,A.R., Shue,Y.L., Lyons,L.A., O'Brien,S.J., Laughlin,T.F.,
Bowling,A.T. and Murray,J.D.
A comparative gene map of the horse (EQUUS caballus)
Genome Res. 9 (12), 1239-1249 (1999)
MEDLINE
20082971
PUBMED
10613847
REFERENCE
2 (bases 1 to 266)
CAetano,A.R.
Direct Submission
Submitted (25-FEB-1999) Veterinary Genetics Lab., University of
California, Davis, One Shields Ave., Davis, CA 95616, USA
FEATURES
Location/Qualifiers
source
1..266
/organism="EQUUS caballus"
/mol_type="genomic DNA"
/db_xref="taxon:9796"
/chromosome="20"
<1..>266
/gene="EDN1"
<1..>266
/misc_feature
/gene="EDN1"
/note="endothelin-1"
ORIGIN
Query Match 7.2%; Score 166; DB 4; Length 266;
Best Local Similarity 83.0%; Pred. No. 5.4e-27;
Matches 229; Conservative 0; Mismatches 35; Indels 12; Gaps 3;
Qy 233 CCAAAAAGCAGAGAGTGTGGAATTTTCCCAACGAGAAAGAACTCAGGTAGCGAGA 292
```

```
Db 1 CCAAAAAGCAGAGAGTGTGGAATTTTCCCAAGCAGGAAAGAACTCAGGTAGTAGA 60
Qy 293 AACACCTTTGCTTTTCAATCAGTTTAAACAGCTTCTGAACTCTTCTCTATCATGAGTACTG 352
Db 61 AACAC--TTTACTTTCAATCAGTTTACAGCTTCTGCTCTCTCTCTCTCTCTCTCTCTCT 118
Qy 353 CCTTCTGTTTTAGAGAGACTAAACAGAGACATTGAAAGTCAGGTAAGCTGAATATAAC 412
Db 119 CCTTCTGTTTAAAGAGATTACAGAGAGCTGGAGAG-----CACCCTGAGTTACC 170
Qy 413 ATTGCTGAATGTTTTTCC--TTGTGTATTTTAAACAGGGCTGAGACATTTATCGAGAAAG 470
Db 171 ATTGCTGAATGTTTTTCCCTATGTGTATTTTAAACAGGGACCAAGACACTATCGAGAAAG 230
Qy 471 ACTGGATATATCAAGAAAGAAAGACTGTTCCA 506
Db 231 GCTGGAATAACCAAGAAAGAAAGAAAGACTGTTCCA 266

RESULT 13
BD094780 636 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION
The cell having the potentiality of differentiation into
cardiomyocytes.
ACCESSION
BD094780
VERSION
BD094780.1 GI:22640368
KEYWORDS
WO 0148150-A/47;
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 636)
Umezawa,A., Hata,J., Fukuda,K., Ogawa,S., Sakurada,K., Gojo,S. and
Yamada,Y.
The cell having the potentiality of differentiation into
Patent: WO 0148150-A 47 05-JUL-2001;
JOURNAL
KYOWA HAKKO KOGYO CO LTD,AKIHIRO UMEZAWA,JUNICHI HATA, KEIICHI
FUKUDA, SATOSHI OGAWA,KAZUHIRO SAKURADA,SATOSHI GOJO,YOJI YAMADA
OS Homo sapiens (human)
COMMENT
PN WO 0148150-A/47
PD 05-JUL-2001
PF 02-NOV-2000 WO 2000JP007741
PR 28-DEC-1999 JP 99P 372826, 28-FEB-2000 WO PCTJP0001448 PI
AKIHIRO UMEZAWA,JUNICHI HATA,KEIICHI FUKUDA,SATOSHI OGAWA, PI
KAZUHIRO SAKURADA,SATOSHI GOJO,YOJI YAMADA
PC C12N5/06,C12N5/10,C12N15/09,A61K31/203,A61K35/28,A61K38/19, PC
A61K38/39,
PC A61K38/45,A61K46/00,A61P9/10,A61P41/00,C07K16/28,C12P21/08, PC
C12Q1/02,
PC C12Q1/48,G01N33/577
CC The cell having the potentiality of differentiation into CC
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
source
1..636
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 6.8%; Score 157.4; DB 6; Length 636;
Best Local Similarity 96.4%; Pred. No. 4.7e-25;
Matches 161; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 128 GCACGTTGTTCCGTTATGGACTTCGAAAGCCCTAGGTCCTCAAGAGAGCCTTGAGAAATTTACT 187
Db 234 GCACGTTGTTCCGTTATGGACTTCGAAAGCCCTAGGTCCTCAAGAGAGCCTTGAGAAATTTACT 293
Qy 188 TCCCAACAAGGCAACAGACCGTCAATATAGATGCCAATGTGCTAGCCAAAAGACAGAA 247
Db 294 TCCCAACAAGGCAACAGACCGTCAATATAGATGCCAATGTGCTAGCCAAAAGACAGAA 353
```

```
QY 248 GTGCTGGAAATTTTCCCAAGCAGGAAAGAACTCAGGTGACGAGAA 294
DB 354 GTGCTGGAAATTTTCCCAAGCAGGAAAGAACTCAGGTGACGAGAA 400

RESULT 14
LOCUS BD096322 636 bp DNA linear PAT 27-AUG-2002
DEFINITION Cells capable of differentiating into myocardial cells.
ACCESSION BD096322
VERSION BD096322.1 GI:22641910
KEYWORDS WO 0148151-A/47.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 636)
AUTHORS Umezawa,A., Hata,J., Fukuda,K., Ogawa,S., Sakurada,K., Gojo,S. and
Yamada,Y.
TITLE Cells capable of differentiating into myocardial cells
JOURNAL Patent: WO 0148151-A 47 05-JUL-2001;
KYOWA HAKKO KOGYO CO LTD
COMMENT OS Homo sapiens (human)
PN WO 0148151-A/47
PD 05-JUL-2001
PF 27-DEC-2000 WO 2000JP009323
PR 28-DEC-1999 JP 99P 372826,28-FEB-2000 WO PCTJP0001148 PR
O2-NOV-2000 WO PCTJP0007741
PI AKIHIRO UMEZAWA, JUNICHI HATA, KEIICHI FUKUDA, SATOSHI OGAWA, PI
KAZUHIRO SAKURADA, SATOSHI GOJO, YOJI YAMADA
PC C12N5/06, C12N5/08, C12P21/08, C12Q1/02, A61K35/28, A61K35/44, A61P9/ PC
06, A61P9/04//A61K38/18, C12N15/12
CC Cells capable of differentiating into myocardial cells FH
Key Location/Qualifiers
FT CDS (1)..(639).

FEATURES
source
1..636
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 6.8%; Score 157.4; DB 6; Length 636;
Best Local Similarity 96.4%; Pred. No. 4.7e-25;
Matches 161; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 128 GCACGTTGTTCCGTATGGACTTGGAGCCCTAGGTCCAGAGAGCCCTTGGAGAAATTACT 187
DB 234 GCACGTTGTTCCGTATGGACTTGGAGCCCTAGGTCCAGAGAGCCCTTGGAGAAATTACT 293
QY 188 TCCCAACAAAGGCAACAGACCGGTGAGATGCCAATGTCTAGCCAAAAGACAAGAA 247
DB 294 TCCCAACAAAGGCAACAGACCGGTGAGATGCCAATGTCTAGCCAAAAGACAAGAA 353
QY 248 GTGCTGGAAATTTTCCCAAGCAGGAAAGAACTCAGGTGACGAGAA 294
DB 354 GTGCTGGAAATTTTCCCAAGCAGGAAAGAACTCAGGTGACGAGAA 400

RESULT 15
LOCUS HSET 1167 bp mRNA linear PRI 23-MAR-1995
DEFINITION Human mRNA for endothelin.
ACCESSION Y00749
VERSION Y00749.1 GI:31254
KEYWORDS endothelin; vasoconstrictor peptide.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1167)

AUTHORS Itoh, Y., Yanagisawa, M., Ohkubo, S., Kimura, C., Kosaka, T., Inoue, A.,
Ishida, N., Mitsui, Y., Onda, H., Fujino, M. and Masaki, T.,
Cloning and sequence analysis of cDNA encoding the precursor of a
human endothelium-derived vasoconstrictor peptide, endothelin:
identity of human and porcine endothelin
FEBS Lett. 231 (2), 440-444 (1988)

JOURNAL MEDLINE 88196441
PUBMED 3282927
REFERENCE 2 (bases 1 to 1167)
AUTHORS Itoh, Y.
TITLE Direct Submission
JOURNAL Submitted (05-APR-1988) Itoh Y., Wadai 7, Tsukuba, Ibaraki 300-42,
Japan
COMMENT Data kindly reviewed (01-JUN-1988) by Itoh Y.
FEATURES
source
1..1167
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="pHET4-3"
/tissue_type="placenta"
253..891
/note="unnamed protein product; preproendothelin (AA -17
to 195)"
/codon_start=1
/protein_id="CAA68718.1"
/db_xref="GI:31255"
/db_xref="GOA:P05305"
/db_xref="SWISS-PROT:P05305"
/translation="MDYLLMFSLIPVACQGNPETAVLGAEISAVGENGEGKPTSPSP
WLRNRKSCSSLMDCBVFCHLDIIVNPIPERVPIGLGSPRSKRALENLLPTKA
TDRENRCCASQDKKCNWFQAGKELRAEDIMEKDNWNNHKKGDKCKLGGKCIYQQL
VGRKIRRSSEELRQTRSETNRNSVKSFFHDPKLGKPSERYVTHRAHW"
253..303
/note="put. signal peptide (AA -17 to -1)"
409..471
/product="mature endothelin (AA 36 - 56)"

sig_peptide
mat_peptide

ORIGIN
Query Match 6.8%; Score 157.4; DB 9; Length 1167;
Best Local Similarity 96.4%; Pred. No. 4.5e-25;
Matches 161; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 128 GCACGTTGTTCCGTATGGACTTGGAGCCCTAGGTCCAGAGAGCCCTTGGAGAAATTACT 187
DB 486 GCACGTTGTTCCGTATGGACTTGGAGCCCTAGGTCCAGAGAGCCCTTGGAGAAATTACT 545
QY 188 TCCCAACAAAGGCAACAGACCGGTGAGATGCCAATGTCTAGCCAAAAGACAAGAA 247
DB 546 TCCCAACAAAGGCAACAGACCGGTGAGATGCCAATGTCTAGCCAAAAGACAAGAA 605
QY 248 GTGCTGGAAATTTTCCCAAGCAGGAAAGAACTCAGGTGACGAGAA 294
DB 606 GTGCTGGAAATTTTCCCAAGCAGGAAAGAACTCAGGTGACGAGAA 652

Search completed: July 16, 2004, 20:03:52
Job time : 8815 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2004, 14:43:24 ; Search time 869 Seconds
(without alignments)
11248.681 Million cell updates/sec

Title: US-10-043-715-1_COPY_157700_160000

Perfect score: 2301

Sequence: 1 ctgaatgatctcccaagt.....tagctccacggagagccct 2301

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 337863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: N Geneseq 29Jan04:*
- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002s:*
- 7: Geneseqn2003as:*
- 8: Geneseqn2003bs:*
- 9: Geneseqn2003cs:*
- 10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2301	100.0	186510	9	ADE24797 Human end
2	2241.6	97.4	12461	2	AAX32367 Human end
3	2241.6	97.4	12461	3	AAX35167 Human end
4	2241.6	97.4	12461	3	AAX35167 Human end
5	2241.6	97.4	12461	6	Aaf21289 Human low
6	2241.6	97.4	12461	6	ABK83778 Human CDN
7	2241.6	97.4	12461	7	ABZ96983 Human nuc
8	2241.6	97.4	12461	7	ACA64929 Human EDN
9	2241.6	97.4	14879	3	AAX35170 Human ade
10	2241.6	97.4	14879	3	Aaf21292 Human low
11	2241.6	97.4	14879	7	ABZ96986 Human nuc
12	159	6.9	1250	6	AAD28049 Human end
13	157.4	6.8	636	4	ABV76130 Human end
14	157.4	6.8	636	4	AAX44397 Human ET1
15	157.4	6.8	1167	3	AAX35169 Human ade
16	157.4	6.8	1167	3	Aaf21291 Human low
17	157.4	6.8	1167	3	AAX54135 Preproend
18	157.4	6.8	1167	3	ABZ96985 Human nuc
19	157.4	6.8	1175	1	AAN90770 Human end
20	157.4	6.8	1175	1	AAN90721 Human end
21	157.4	6.8	1251	3	AAX35168 Human ade
22	157.4	6.8	1251	3	Aaf21290 Human low
23	157.4	6.8	1251	6	ABZ34893 Human gen

24	157.4	6.8	1251	6	ABV94467 Breast ca
25	157.4	6.8	1251	6	ABV78072 Hypoxia-r
26	157.4	6.8	1251	7	ABZ96984 Human nuc
27	157.4	6.8	1251	7	ABX08802 Angiogene
28	156.4	6.8	566	6	ABV94649 Human pan
29	155.8	6.8	512	8	ACH33801 Human end
30	155.8	6.8	577	7	ABZ84592 Toxicolog
31	101.4	4.4	695	6	ABT09430 Phase-1 R
32	100.8	4.4	1838	1	AAN90769 cDNA of e
33	99.2	4.3	749	1	AAN90720 Porcine e
34	97	4.2	300	6	ABL72947 Corn tass
35	60	2.6	60	6	ABN41499 Human spl
36	52.2	2.3	2000	7	ADA71938 Rice gene
37	48.6	2.1	6247	6	ABK39923 Human che
38	48.6	2.1	11371	6	ABL32773 Human imm
39	47.4	2.1	16287	6	ABL32672 Human imm
40	47	2.0	6301	6	ABN80101 Human che
41	47	2.0	33053	6	ABQ67006 Human ang
42	46.8	2.0	6956	6	ABL70225 Chemicall
43	46	2.0	5397	6	ABL33045 Human imm
44	45.6	2.0	14542	6	ABK31234 Signal tr
45	45.6	2.0	14542	6	ABL70191 Chemicall

ALIGNMENTS

RESULT 1
ADE24797
ID ADE24797 standard; DNA; 186510 BP.
XX ()
AC ADE24797;
DT 29-JAN-2004 (first entry)
DE Human endothelin-1, EDN1, gene.
KW ds; gene; human; vascular disease; endothelin-1; EDN1;
KW coronary artery disease; myocardial infarction.
XX Homo sapiens.
FH Key Location/Qualifiers
FT variation replace(157790,C)
FT /tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(159908,T)
FT /tag= b
FT /standard_name= "Single nucleotide polymorphism"
XX US2003143544-A1.
PN 31-JUL-2003.
XX 09-JAN-2002; 2002US-00043715.
XX 09-JAN-2002; 2002US-00043715.
XX (VITI-) VITIVITY INC.
XX McCarthy J;
XX WPI; 2003-874790/81.
XX P-PSDB; ADE24798, ADE24801.
XX Identifying a subject as a candidate for a particular therapy to treat a
XX vascular disease or disorder, particularly coronary artery disease or
XX myocardial infarction, comprises detecting polymorphisms of the
XX endothelin-1 gene.
XX Claim 58; SEQ ID NO 1; 177pp; English.
XX The invention relates to a method of identifying a subject as a candidate

CC for a particular therapy to treat a vascular disease or disorder
CC comprises determining the presence of nucleotides at polymorphic regions
CC of an endothelin-1 (EDN1) gene. The invention is used to determine risk
CC of developing a vascular disease or disorder, particularly coronary
CC artery disease or myocardial infarction. The present sequence represents
CC the human endothelin-1 gene.

XX
SQ Sequence 186510 BP; 56192 A; 36427 C; 37117 G; 56774 T; 0 U; 0 Other;

Query Match	100.0%;	Score 2301;	DB 9;	Length 186510;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2301;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	CTGAATGATGCTCCAAAGTCTATGTGATGAGTCTCTGTGTGCTGCCAGTGGAAATAGGTG 60			
DB 157700	CTGAATGATGCTCCAAAGTCTATGTGATGAGTCTCTGTGTGCTGCCAGTGGAAATAGGTG 157759			
QY 61	TGTCATGTCGTCATTTAAAGACTATTAATACACTAATATAGTTTCTTCTCTCTTTGG 120			
DB 157760	TGTCATGTCGTCATTTAAAGACTATTAATACACTAATATAGTTTCTTCTCTCTTTGG 157819			
QY 121	ATAATAGGACAGTTCTTCCGTATGGACTTGAAGCCCTAGGTCCAAGAGGCTTTGGAGA 180			
DB 157820	ATAATAGGACAGTTCTTCCGTATGGACTTGAAGCCCTAGGTCCAAGAGGCTTTGGAGA 157879			
QY 181	ATTACTTCCCAAGGACACAGCCGTGAATATAGTATGCTAGTCCAAAG 240			
DB 157880	ATTACTTCCCAAGGACACAGCCGTGAATATAGTATGCTAGTCCAAAG 157939			
QY 241	ACAGAAGTGTGGAAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAGAAACACCTT 300			
DB 157940	ACAGAAGTGTGGAAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAGAAACACCTT 157999			
QY 301	TGCTTTTCAATCAGTTTAAAGCCCTCTGAATCTCTCTATCATGCTAGTCCCTTCTG 360			
DB 158000	TGCTTTTCAATCAGTTTAAAGCCCTCTGAATCTCTCTATCATGCTAGTCCCTTCTG 158059			
QY 361	TTTTAGAGAGACTAACAGAGACATTTGAAGTCAGGTAAGCTGAATATACATTTGCTGA 420			
DB 158060	TTTTAGAGAGACTAACAGAGACATTTGAAGTCAGGTAAGCTGAATATACATTTGCTGA 158119			
QY 421	AATGTTTTTCCCTGTGATTTTAAACAGGCTTGAAGACATTTATGAGAAAGACTGGAATAA 480			
DB 158120	AATGTTTTTCCCTGTGATTTTAAACAGGCTTGAAGACATTTATGAGAAAGACTGGAATAA 158179			
QY 481	TCATAAGAAAGAAAGACTGTTCAGCTTGGGAAAGTGTATTTATCAGCAGTTAGT 540			
DB 158180	TCATAAGAAAGAAAGACTGTTCAGCTTGGGAAAGTGTATTTATCAGCAGTTAGT 158239			
QY 541	GAGAGAAAGAAATCAGAAAGTTCAGAGAACACCTTAAGACAAACACGATTAAGAGGG 600			
DB 158240	GAGAGAAAGAAATCAGAAAGTTCAGAGAACACCTTAAGACAAACACGATTAAGAGGG 158299			
QY 601	AGGAGAAAGAAATTAGGTAAGGTTTCAAGACAACTAGCCCAAGTCAGTATGCCAG 660			
DB 158300	AAGGAGAAAGAAATTAGGTAAGGTTTCAAGACAACTAGCCCAAGTCAGTATGCCAG 158359			
QY 661	CAGCTGTCTTCCAGCCCTTCTTACCAGGCTGAAAGACTTTAGAAACACAGTAGCAG 720			
DB 158360	CAGCTGTCTTCCAGCCCTTCTTACCAGGCTGAAAGACTTTAGAAACACAGTAGCAG 158419			
QY 721	AGGAGATCTATGCTATATAGATTTAAAGAGGCAAGATCCCTCTTAATATTTCCA 780			
DB 158420	AGGAGATCTATGCTATATAGATTTAAAGAGGCAAGATCCCTCTTAATATTTCCA 158479			
QY 781	TGAAGCTCTGGAATGCAAAACCATGCTCTGTACTTTTAGCACATACCATTTCTATC 840			
DB 158480	TGAAGCTCTGGAATGCAAAACCATGCTCTGTACTTTTAGCACATACCATTTCTATC 158539			
QY 841	AGGTAGATTTTCCAAACCAAAATATATCCAGAGATCCCTTTGTTCATTTGGTTATATAC 900			
DB 158540	AGGTAGATTTTCCAAACCAAAATATATCCAGAGATCCCTTTGTTCATTTGGTTATATAC 158599			

QY 901	CTTTGCTCTCTGAGTCAATGATTTTACCACTTTCCCTGAGAAATCGAAATCATTTTGG 960			
DB 158600	CTTTGCTCTCTGAGTCAATGATTTTACCACTTTCCCTGAGAAATCGAAATCATTTTGG 158659			
QY 961	GGAGGGGACATTTAGAAAAAGAAATCAAAAGTGTGATGGATAATCAAAATCTTCAATAAGTT 1020			
DB 158660	GGAGGGGACATTTAGAAAAAGAAATCAAAAGTGTGATGGATAATCAAAATCTTCAATAAGTT 158719			
QY 1021	GCAGTATTTCAGATGGCCAAAGAAAAATAAAGTCATTAGATAGGTTGGTAGAATTTAG 1080			
DB 158720	GCAGTATTTCAGATGGCCAAAGAAAAATAAAGTCATTAGATAGGTTGGTAGAATTTAG 158779			
QY 1081	ACATGCTGTTTTTTCAGGTTTATGCTCTTTTTTTTTTTTTTTTTTTTTTAAATAGGAAATG 1140			
DB 158780	ACATGCTGTTTTTTCAGGTTTATGCTCTTTTTTTTTTTTTTTTTTTTTTAAATAGGAAATG 158839			
QY 1141	TGTTTTGGTGCAGAGCAATGTCAATCCAAAGAGCTCTCTTTTTCTCGTCAGTCAATGTG 1200			
DB 158840	TGTTTTGGTGCAGAGCAATGTCAATCCAAAGAGCTCTCTTTTTCTCGTCAGTCAATGTG 158899			
QY 1201	CTGGGACAGAGAGGATCTGGATTAGGCAACATCATAGAGTTGCTCTGAGTCTCTTT 1260			
DB 158900	CTGGGACAGAGAGGATCTGGATTAGGCAACATCATAGAGTTGCTCTGAGTCTCTTT 158959			
QY 1261	GGTGATAACCCCTTCCAAATCTTAAACTTTTGGAAATTCACAAAGCTCAAAGGAGGAAACCT 1320			
DB 158960	GGTGATAACCCCTTCCAAATCTTAAACTTTTGGAAATTCACAAAGCTCAAAGGAGGAAACCT 159019			
QY 1321	ACTCTGATCTACACATGTTCTGCATTTTCTATCATGCTCTATGGAATCTCTCTTA 1380			
DB 159020	ACTCTGATCTACACATGTTCTGCATTTTCTATCATGCTCTATGGAATCTCTCTTA 159079			
QY 1381	GAATCCAGTGGCAAGAGTCTTATGATTAAGTGTCTGAGCTCAGGCCAGGAGTCAAT 1440			
DB 159080	GAATCCAGTGGCAAGAGTCTTATGATTAAGTGTCTGAGCTCAGGCCAGGAGTCAAT 159139			
QY 1441	GAACTACTTCTGAGTTATTTACTGATTTGTTGGGAGCCCTCAGCTATCGTTTCTTTC 1500			
DB 159140	GAACTACTTCTGAGTTATTTACTGATTTGTTGGGAGCCCTCAGCTATCGTTTCTTTC 159199			
QY 1501	ACACCTGCTTATGAGAGTATCCATATTTATGTTCCAGCCAGTAAATGCTCCCAACGAGA 1560			
DB 159200	ACACCTGCTTATGAGAGTATCCATATTTATGTTCCAGCCAGTAAATGCTCCCAACGAGA 159259			
QY 1561	TCAGTTTCTGAACTAACCTGGAATTTTATGTTTATGTTTATGTTTATGTTTATGTTTAT 1620			
DB 159260	TCAGTTTCTGAACTAACCTGGAATTTTATGTTTATGTTTATGTTTATGTTTATGTTTAT 159319			
QY 1621	ACATTCAGTTCTTCCCTCTGATTTCTCCTCTAAACATATTAGGCTGCAAAAAAATAA 1680			
DB 159320	ACATTCAGTTCTTCCCTCTGATTTCTCCTCTAAACATATTAGGCTGCAAAAAAATAA 159379			
QY 1681	ATCTTTTAAAAAATTAATGCCATAAAGTATTTGCTCTGGGCTACTGATGCTCTTTTC 1740			
DB 159380	ATCTTTTAAAAAATTAATGCCATAAAGTATTTGCTCTGGGCTACTGATGCTCTTTTC 159439			
QY 1741	TTTTTCTCTCTTTTCACTAAGTCCAGTCAATTTTATTAAGATGCCATTAATCTTCAAA 1800			
DB 159440	TTTTTCTCTCTTTTCACTAAGTCCAGTCAATTTTATTAAGATGCCATTAATCTTCAAA 159499			
QY 1801	ACCTATGCTGAGTTCTCCTCAAGCAGGTCACATAGTATGATGAAGTTGGGATGGGCTAGC 1860			
DB 159500	ACCTATGCTGAGTTCTCCTCAAGCAGGTCACATAGTATGATGAAGTTGGGATGGGCTAGC 159559			
QY 1861	GAAGAAACAGAAACAACTCTAGTTTATTTAAACCTGATTTTACTGCCACCTTCCCTTA 1920			
DB 159560	GAAGAAACAGAAACAACTCTAGTTTATTTAAACCTGATTTTACTGCCACCTTCCCTTA 159619			
QY 1921	GACTTCACCATATGACCCCTCGCTCCCATTTCTAAGCATAGGGGAGGCTTTATTTTACA 1980			
DB 159620	GACTTCACCATATGACCCCTCGCTCCCATTTCTAAGCATAGGGGAGGCTTTATTTTACA 159679			
QY 1981	ATCGTAATAGATATCACTTGAGGTTTTTATCAAGAGTTCGCGGGGTGTAAGTTTAC 2040			

Db 159680 ATGGTAATAGATATCACCTTGAGTTTATCAAGAGTTTGGCGGGTGGTGAAGTTTCA 159739
Qy 2041 AACCGATTCCAGTTTGTGTCAGATCTAATTTTACATGTTCTTTTGGCAAGG 2100
Db 159740 AACCGATTCCAGTTTGTGTCAGATCTAATTTTACATGTTCTTTTGGCAAGG 159799
Qy 2101 GTGATTTTAAATAACATTTGTTTCTCTTATCTTGTCTTTTATAGTTCGAGACCAT 2160
Db 159800 GTGATTTTAAATAACATTTGTTTCTCTTATCTTGTCTTTTATAGTTCGAGACCAT 159859
Qy 2161 GAGAAACAGCGTCAATCATCTTTTCATGATCCCAAGCTGAAGCAAGCCCTCCAGAGA 2220
Db 159860 GAGAAACAGCGTCAATCATCTTTTCATGATCCCAAGCTGAAGCAAGCCCTCCAGAGA 159919
Qy 2221 GGGTTATGTGACCCCAACAGCAGACATTTGTTGACAGACCTTCGGGGCTCTCTGAAGCC 2280
Db 159920 GGGTTATGTGACCCCAACAGCAGACATTTGTTGACAGACCTTCGGGGCTCTCTGAAGCC 159979
Qy 2281 ATAGCCTCCAGGAGAGCCCT 2301
Db 159980 ATAGCCTCCAGGAGAGCCCT 160000

RESULT 2
ID AAX32367
XX AAX32367 standard; DNA; 12461 BP.
XX AC AAX32367;
XX DT 16-JUN-1999 (first entry)
XX DE Human endothelin-1 (ET-1) gene sequence.
XX KW Pulmonary hypertension; therapeutic; aerosolized; endothelin-1; ET-1;
XX LW lung; antisense; ss.
XX OS Homo sapiens.
XX PN WO9911778-Al.
XX PD 11-MAR-1999.
XX PF 02-SEP-1998; 98WO-GB002584.
XX PR 02-SEP-1997; 97GB-00018487.
XX PS (UYSH-) UNIV SHEFFIELD.
XX PA Higenbottam T, McCormack K, Smith A;
XX PI New composition containing an aerosolized antisense ET-1 molecule -
XX PT useful for treating pulmonary hypertension.
XX PS Disclosure; Fig.1; 37pp; English.
XX CC The invention relates to a method for treating pulmonary hypertension by
XX delivering a therapeutic composition, comprising an aerosolized antisense
XX endothelin-1 (ET-1) molecule, to the lungs of a patient. The composition
XX can be used in a method for determining the efficacy of the treatment for
XX e.g. when studying molecules and observing the effects of the composition
XX on an animal model system hypersensitive to antisense ET-1. The method is
XX useful for treating pulmonary hypertension. The aerosolized antisense ET-
XX 1 molecule permits inhibition of the ET-1 transcription, which relieves
XX pulmonary hypertension. Its use avoids side effects caused by alternative
XX therapies. The present sequence represents the genomic DNA sequence of
XX human ET-1 gene
SQ Sequence 12461 BP; 3372 A; 2728 C; 2816 G; 3544 T; 0 U; 1 Other;
Query Match 97.4%; Score 2241.6; DB 2; Length 12461;

Best Local Similarity 99.3%; Pred. No. 0;
Matches 2294; Conservative 1; Mismatches 5; Indels 10; Gaps 4;
Qy 1 CTGAAATGATGCTCCCAAGTGCTATGCTGATGAGCTCTTGTGTCGCCAGTGAATAGG 60
Db 7056 CTGAAATGATGCTCCCAAGTGCTATGCTGATGAGCTCTTGTGTCGCCAGTGAATAGG 7115
Qy 61 TGTCCATGTCTCATTTTAAAGACTATTAATACACTAATATAGTTTCTTCTCTCTTGG 120
Db 7116 TGTCCATGTCTCATTTTAAAGACTATTAATACACTAATATAGTTTCTTCTCTCTTGG 7175
Qy 121 ATATAGGACAGTGTTCGTTGATGAGCTTGGAGCCCTAGGTCCAAAGAGAGCCCTTGGAGA 180
Db 7176 ATATAGGACAGTGTTCGTTGATGAGCTTGGAGCCCTAGGTCCAAAGAGAGCCCTTGGAGA 7235
Qy 181 ATTTACTTCCCAAAAGGCAACAGACCGTGAAAAATAGATGCCAATGTCTAGTCCAAAAAG 240
Db 7236 ATTTACTTCCCAAAAGGCAACAGACCGTGAAAAATAGATGCCAATGTCTAGTCCAAAAAG 7295
Qy 241 ACAAGAAGTCTGGAAATTTTTCGCAAGCAGGAAAAAGAACTCAGGTGAGCAGAAAAACCTT 300
Db 7296 ACAAGAAGTCTGGAAATTTTTCGCAAGCAGGAAAAAGAACTCAGGTGAGCAGAAAAACCTT 7355
Qy 301 TGCCTTTCAATCAGTTTAAAGCCTCTGAACTCCTTCTCTATCATGTGCTACTGCTTCCTG 360
Db 7356 TGCCTTTCAATCAGTTTAAAGCCTCTGAACTCCTTCTCTATCATGTGCTACTGCTTCCTG 7415
Qy 361 TTTTAGAGAGACTAACAGAGACATTGAAAGTCAAGGCTAAAGCTGAATATAAATGCTGTA 420
Db 7416 TTTTAGAGAGACTAACAGAGACATTGAAAGTCAAGGCTAAAGCTGAATATAAATGCTGTA 7475
Qy 421 AATGTTTTTCTTGTGTATTTTAAAGGCTGAAGACATTTGGAGAAAGACTGGAAATAA 480
Db 7476 AATGTTTTTCTTGTGTATTTTAAAGGCTGAAGACATTTGGAGAAAGACTGGAAATAA 7535
Qy 481 TCATAAGAAAGGAAAGACTGTTTCCAAAGCTTGGGAAAAAGTGTATTATCAGCAGTTAGT 540
Db 7536 TCATAAGAAAGGAAAGACTGTTTCCAAAGCTTGGGAAAAAGTGTATTATCAGCAGTTAGT 7595
Qy 541 GAGAGAGAAAGAAATCAGAAAGAGTTTCAAGAGCAACCTAAGACAAACAGGTGAAGAGG 600
Db 7596 GAGAGAGAAAGAAATCAGAAAGAGTTTCAAGAGCAACCTAAGACAAACAGGTGAAGAGG 7655
Qy 601 AAGGAGAGAAAGAAATTAGTAAAGAGTTTCAAGAGCAACCTAAGACCACTGATGCCAG 660
Db 7656 AAGGAGAGAAAGAAATTAGTAAAGAGTTTCAAGAGCAACCTAAGACCACTGATGCCAG 7715
Qy 661 CAGCCTGTTCTCCTCAGCCCTTCTTACCAGGAGGTGAAGAGCTTAGAAAAACAGTAGCAG 720
Db 7716 CAGCCTGTTCTCCTCAGCCCTTCTTACCAGGAGGTGAAGAGCTTAGAAAAACAGTAGCAG 7775
Qy 721 AGGAGATCTATGATCTCTATAGATTTAAAGGAGCAAGAAATCCCTCTTAAATATTTCCA 780
Db 7776 AGGAGATCTATGATCTCTATAGATTTAAAGGAGCAAGAAATCCCTCTTAAATATTTCCA 7835
Qy 781 TGAAGCTCTGGAATGCAAAACCGATGCTCTGTACTTTTAGCACATACCATTTCATCTAC 840
Db 7836 TGAAGCTCTGGAATGCAAAACCGATGCTCTGTACTTTTAGCACATACCATTTCATCTAC 7895
Qy 841 AGTGAATTTCCCAACCAAAATATATCCAGAGATGCTTTGTTCATTTGGGTATATACAGC 900
Db 7896 AGTGAATTTCCCAACCAAAATATATCCAGAGATGCTTTGTTCATTTGGGTATATACAGC 7955
Qy 901 CTTTGGCTCTCTGAGTCAATGTATTTTACCACCTTTCCCTCAGAAATCGAAAAATCATTTGG 960
Db 7956 CTTTGGCTCTCTGAGTCAATGTATTTTACCACCTTTCCCTCAGAAATCGAAAAATCATTTGG 8015
Qy 961 GGAGCGGACATTTAGAAAAAGAAATCAAGTGTCAATGATGAATCAAAATTTCTCAATAAGTT 1020
Db 8016 GGAGCGGACATTTAGAAAAAGAAATCAAGTGTCAATGATGAATCAAAATTTCTCAATAAGTT 8075
Qy 1021 GCAGTTATTTCAGATGGCCAAAGGAAAAATAAAGTCAATGATAGGTTCGTAGAAATTTAG 1080

Db 8076 GCAGTTATTAGATGGCCAAAGGAAAAATAAAGTCATTAGATAGGGTTGGTAGAATTTAG 8135
QY 1081 AACATGCTGTTTTCAGGTTTATGGTC-----TTTTTTTTTTTTTTTTTTTTTTTTTAATAGGG 1135
Db 8136 AACATGCTGTTTTCAGGTTTATGGTC-----TTTTTTTTTTTTTTTTTTTTTTTTTAATAGGG 8195
QY 1136 AAATGTTTGGTGAGAGCCATGTCTATTCACAAAAGCTCTCTTTTCTGTCAGTC 1195
Db 8196 AAATGTTTGGTGAGAGCCATGTCTATTCACAAAAGCTCTCTTTTCTGTCAGTC 8255
QY 1196 ATGTCCTGGGACAGAGAGGGATCTGGATTAGGCAACATCATAGATTGCTCTCAGCTGC 1255
Db 8256 ATGTCCTGGGACAGAGAGGGATCTGGATTAGGCAACATCATAGATTGCTCTCAGCTGC 8315
QY 1256 TCTTTGGTGATAACCCCTTCCAAATCCTAAACCTTTTGGAAATCAAGCTCAAAAGGAGGA 1315
Db 8316 TCTTTGGTGATAACCCCTTCCAAATCCTAAACCTTTTGGAAATCAAGCTCAAAAGGAGGA 8375
QY 1316 AACCTACTCTGTACTACCACTGCTCTGCAATTTTCTATCATGGTCTATGGAACCTTC 1375
Db 8376 AACCTACTCTGTACTACCACTGCTCTGCAATTTTCTATCATGGTCTATGGAACCTTC 8435
QY 1376 TCTTAGAATCCAGTGGCAAGAGTTCTATGATTAAGTGTCTGAGCTCAGGCCAGGCA 1435
Db 8436 TCTTAGAATCCAGTGGCAAGAGTTCTATGATTAAGTGTCTGAGCTCAGGCCAGGCA 8495
QY 1436 GTCATGAATCTCTGAGTTATTTACTACTGATTTTGGGGAGCCTCAGCTATCGTT 1495
Db 8496 GTCATGAATCTCTGAGTTATTTACTACTGATTTTGGGGAGCCTCAGCTATCGTT 8555
QY 1496 TCTTCACACCTGCTATGAGAGTATCCATATTTATGTCGAGCGCAGTATGCTCCCCA 1555
Db 8556 TCTTCACACCTGCTATGAGAGTATCCATATTTATGTCGAGCGCAGTATGCTCCCCA 8614
QY 1556 CGAGATCAGTTCTGAACTAACCTGGAAATTTTATGAGGTTTATTAAGCAACTATTA 1615
Db 8615 CGAGATCAGTTCTGAACTAACCTGGAAATTTTATGAGGTTTATTAAGCAACTATTA 8674
QY 1616 AATCAACATTCAGTTCTCCCTCTGATTTCTCTCTGTAACCAATTAGGCTCGAAAAA 1675
Db 8675 AATCAACATTCAGTTCTCCCTCTGATTTCTCTCTGTAACCAATTAGGCTCGAAAAA 8734
QY 1676 AAAAAATCTTTTAAAAATAATGTCATAAAGTATTTGCTCTGGCCCTACTGATGCTTC 1735
Db 8735 AAAAAATCTTTTAAAAATAATGTCATAAAGTATTTGCTCTGGCCCTACTGATGCTTC 8794
QY 1736 TTTTCTTTTCTCTCTTTTCACTAAGTCACCGTCATTTTATTAAGATGCGCATAACTAT 1795
Db 8795 TTTTCTTTTCTCTCTTTTCACTAAGTCACCGTCATTTTATTAAGATGCGCATAACTAT 8854
QY 1796 TCAAAACCTATGCTGAGTTCTCCTCAGGAGGCTCAGATGATCAAGGTTGGGATGGG 1855
Db 8855 TCAAAACCTATGCTGAGTTCTCCTCAGGAGGCTCAGATGATCAAGGTTGGGATGGG 8914
QY 1856 CTACCGAGAGACACAGAACTCTAGTTTATTTTAAACCTGTATTTACTGCCCACTTCC 1915
Db 8915 CTACCGAGAGACACAGAACTCTAGTTTATTTTAAACCTGTATTTACTGCCCACTTCC 8974
QY 1916 CTTAGACTTGACCATATGACCCCTCGCT-CCCATCTTAAGCATAGGGGAGGCTTTATT 1974
Db 8975 CTTAGACTTGACCATATGACCCCTCGCT-CCCATCTTAAGCATAGGGGAGGCTTTATT 9034
QY 1975 TTTAATGTTAATA- -GATATCATCTGAGTTTATCAAGAGTTGCGCGGGGCTG 2031
Db 9035 TTTAATGTTAATAAGATGATATCACTTGGGTTTATCAAGAGTTGCGCGGGGCTG 9094
QY 2032 AAAGTTCAACACGATTCAGGTTTGTGTCGAGATTTCTAAATTTTACATGTTCTTT 2091
Db 9095 AAAGTTCAACACGATTCAGGTTTGTGTCGAGATTTCTAAATTTTACATGTTCTTT 9154
QY 2092 TGCCAAAAGGTTGATTTTTTTTAAATAACATTTGTTTCTCTTATCTGCTTTATAGGTC 2151
Db 9155 TGCCAAAAGGTTGATTTTTTTTAAATAACATTTGTTTCTCTTATCTGCTTTATAGGTC 9214

QY 2152 GGAGACCATGAGAAACAGCGTCAAAATCATCTTTTTCATGATCCCAAGCTGAAAGGCAAGCC 2211
Db 9215 GGAGACCATGAGAAACAGCGTCAAAATCATCTTTTTCATGATCCCAAGCTGAAAGGCAAGCC 9274
QY 2212 CTCCAGAGCGCTTATGTGACCCACACCGAGCAGCATTTGGTGACAGACCTTCGGGGCCCTG 2271
Db 9275 CTCCAGAGCGCTTATGTGACCCACACCGAGCAGCATTTGGTGACAGACCTTCGGGGCCCTG 9334
QY 2272 TCTGAAGCCATAGCCTCCACGGAGAGCCCT 2301
Db 9335 TCTGAAGCCATAGCCTCCACGGAGAGCCCT 9364

RESULT 3

AAA35167

ID AAA35167 standard; DNA; 12461 BP.

XX AAA35167;

XX AC AAA35167;

XX DT 28-JUL-2000 (first entry)

XX XX

DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:41.

XX XX

KW Human; adenosine receptor; low adenosine antisense oligonucleotide;

KW phosphorothioate; impaired respiration; inflammation; allergy;

KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;

KW antiasthmatic; cytosolic; analgesic; impaired airway;

KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;

KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;

KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;

KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.

XX XX

OS Homo sapiens.

XX XX

FN WO200009525-A2.

XX XX

PD 24-FEB-2000.

XX XX

PF 03-AUG-1999; 99WO-US017712.

XX XX

PR 03-AUG-1998; 98US-0095212P.

XX XX

PA (UYEC-) UNIV EAST CAROLINA.

XX XX

PI Nyce JW;

XX XX

WPI; 2000-205971/18.

XX XX

PT New antisense oligonucleotides useful for treating e.g. pulmonary

PT vasoconstriction, inflammation, allergies, asthma, hypertension, or

PT bronchitis, emphysema, respiratory distress syndrome, ischemia or

PT cancers.

XX XX

PS Disclosure; Page 1198-1201; 1343pp; English.

XX XX

CC The present invention describes a new composition comprising an antisense oligonucleotide (ON) with low adenosine (up to 15%), which targets nucleic acids involved in bronchoconstriction, allergies, and/or inflammation. The ON can have antiinflammatory, antiallergic, antiasthmatic, cytosolic and analgesic activities. The compositions are useful for the treatment of diseases associated with inflammation, impaired airways, including lung disease and diseases whose secondary effects afflict the lungs of a subject. They can be used for treating e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma, impaired respiration, respiratory distress syndrome, pain, cystic fibrosis, pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), and cancers such as leukaemias, lymphomas, carcinomas, and cancers which may metastasise to the lungs, including breast and prostate cancer. The reduction of the adenosine content of the ONs reduces side effects. The A-containing ONs break down with the release of deoxyadenosine which activates adenosine receptors causing bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the

CC nucleotide sequences given in the sequence listing from the present invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185 sequences are also called SEQ ID NO:1 to 185, but the sequences differ from the previously named sequences. SEQ ID NO:11 to 1680 (AAA32323 to AAA3392) are specifically claimed. The present invention does not match sequences given in the disclosure of the present invention do not match up with their corresponding SEQ ID NO: sequences given in the sequence listing

XX Sequence 12461 BP; 3372 A; 2728 C; 2816 G; 3544 T; 0 U; 1 Other;

Query Match 97.4%; Score 2241.6; DB 3; Length 12461;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2294; Conservative 1; Mismatches 5; Indels 10; Gaps 4;

QY 1 CTGAAATGATGTCCTCCAAAGTGTATGTCATGAGTCCCTGTGTGCGCCAGTGGAAATAGTG 60
DB 7056 CTGAAATGATGTCCTCCAAAGTGTATGTCATGAGTCCCTGTGTGCGCCAGTGGAAATAGTG 7115
QY 61 TGTCCATGTCATTTTAAAGACTATTAATACACTAATATAGTTTCTTCTCTCTTTGG 120
DB 7116 TGTCCATGTCATTTTAAAGACTATTAATACACTAATATAGTTTCTTCTCTCTTTGG 7175
QY 121 ATAATAGCAGGTGTTCCGATGACCTTGAAGCCCTAGTCCCAAGAGAGCCCTGGAGA 180
DB 7176 ATAATAGCAGGTGTTCCGATGACCTTGAAGCCCTAGTCCCAAGAGAGCCCTGGAGA 7235
QY 181 ATTTACTTCCCAAGGCAACAGACCGTGAATAGATGCTGCTAGCCAAAG 240
DB 7236 ATTTACTTCCCAAGGCAACAGACCGTGAATAGATGCTGCTAGCCAAAG 7295
QY 241 ACAAGAAGTGTGGAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAGAAACACCTT 300
DB 7296 ACAAGAAGTGTGGAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAGAAACACCTT 7355
QY 301 TGCTTTTCAATCAGTTTAAACAGCTTCTGAACTCTCTCTCATGTTGCTGCTCTCTG 360
DB 7356 TGCTTTTCAATCAGTTTAAACAGCTTCTGAACTCTCTCTCATGTTGCTGCTCTCTG 7415
QY 361 TTTTAGAGAGACTAACAGAGACATTTGAAGTCAAGGTAAAGTGAATATATCATTTGCTGA 420
DB 7416 TTTTAGAGAGACTAACAGAGACATTTGAAGTCAAGGTAAAGTGAATATATCATTTGCTGA 7475
QY 421 AATGTTTTTCTTGTTGTTTAAACAGGCTTGAAGACATTTGAGAAAGACTCGAATAA 480
DB 7476 AATGTTTTTCTTGTTGTTTAAACAGGCTTGAAGACATTTGAGAAAGACTCGAATAA 7535
QY 481 TCATAAGAAAGGAAAGACTGTTTCAAGCTTGGGAAAGAGTGTATTTATCAGCAGTTAGT 540
DB 7536 TCATAAGAAAGGAAAGACTGTTTCAAGCTTGGGAAAGAGTGTATTTATCAGCAGTTAGT 7595
QY 541 GAGAGGAGAAAGAAATCAGAGAGTTCAGAGGACCTAGACAAACCCAGGTGAAGGG 600
DB 7596 GAGAGGAGAAAGAAATCAGAGAGTTCAGAGGACCTAGACAAACCCAGGTGAAGGG 7655
QY 601 AAGGAGAGAAAGAAATTAGGTAAGAGTTTCAAGAACTAGCCCCAGTCAAGTGGCCAG 660
DB 7656 AAGGAGAGAAAGAAATTAGGTAAGAGTTTCAAGAACTAGCCCCAGTCAAGTGGCCAG 7715
QY 661 CAGCCTGTTCTCAGCCCTTCTTACCGGGGAGGTGAAGACTTAGAAACAGTAGCAG 720
DB 7716 CAGCCTGTTCTCAGCCCTTCTTACCGGGGAGGTGAAGACTTAGAAACAGTAGCAG 7775
QY 721 AGGAGATCTATGATCTTATGATTTAAAGGAGCAAGAAAGATCCCTCTTAAATATTTCCA 780
DB 7776 AGGAGATCTATGATCTTATGATTTAAAGGAGCAAGAAAGATCCCTCTTAAATATTTCCA 7835
QY 781 TGAAGCTCTGGAATGCAACCGATGCTCTGTAATTTTATGAGACATACATTTCACTAC 840
DB 7836 TGAAGCTCTGGAATGCAACCGATGCTCTGTAATTTTATGAGACATACATTTCACTAC 7895
QY 841 AGGTAGATTTCCCAACCAAAATATATCAGAGATGCTTTGCTAGGTTTATATACAGC 900

DB 7896 AGGTAGATTTCCCAACCAAAATATATCCAGAGATGCTTTGTCTATTTGGTTATATACAGC 7955
QY 901 CTTTGGCTCTCTGAGTCAATATGATTTTACACATTTCCCTGAGAAATCGAAAAATCATTTGG 960
DB 7956 CTTTGGCTCTCTGAGTCAATATGATTTTACACATTTCCCTGAGAAATCGAAAAATCATTTGG 8015
QY 961 GAGCGGCAATTTAGAAAGAAATCAAAAGTGTGATGATAATCAAAATTTTCAATAAGTT 1020
DB 8016 GAGCGGCAATTTAGAAAGAAATCAAAAGTGTGATGATAATCAAAATTTTCAATAAGTT 8075
QY 1021 GCAGTTATTCAGATGGCCAAAGGAAATAAAGTCAATAGATAGGTTGCTGATAGATTTAG 1080
DB 8076 GCAGTTATTCAGATGGCCAAAGGAAATAAAGTCAATAGATAGGTTGCTGATAGATTTAG 8135
QY 1081 AACATGCTGTTTTTCAAGTTTATGTC-----TTTTTTTTTTTTTTTTTTTTTAAATAGGG 1135
DB 8136 AACATGCTGTTTTTCAAGTTTATGTC-----TTTTTTTTTTTTTTTTTTTTTAAATAGGG 8195
QY 1136 AATGTTGTTTGGTGAGAGCCATGTCATCCAAAAGTCTCTCTTTTCTCGGTGAGTC 1195
DB 8196 AATGTTGTTTGGTGAGAGCCATGTCATCCAAAAGTCTCTCTTTTCTCGGTGAGTC 8255
QY 1196 ATGTGCTGGACAGAGAGGATCTGTGATTAGGCAACATCATAGAGTTGCTCTGAGCTGC 1255
DB 8256 ATGTGCTGGACAGAGAGGATCTGTGATTAGGCAACATCATAGAGTTGCTCTGAGCTGC 8315
QY 1256 TCTTGGTGATAACCCCTTCCAAATCTTAACTTTTGGAAATTCACAAAGTCAAGAGGAGA 1315
DB 8316 TCTTGGTGATAACCCCTTCCAAATCTTAACTTTTGGAAATTCACAAAGTCAAGAGGAGA 8375
QY 1316 AACCTACTCTCTGATCTACCAATGTTCTGATTTTCTATCATGTTCTATGGAATCTTC 1375
DB 8376 AACCTACTCTCTGATCTACCAATGTTCTGATTTTCTATCATGTTCTATGGAATCTTC 8435
QY 1376 TCTTAGAAATCCAGTGGCAAGAGTTCTATGATTTAAAGTGTCTGAGCTCAGGCCAGGCA 1435
DB 8436 TCTTAGAAATCCAGTGGCAAGAGTTCTATGATTTAAAGTGTCTGAGCTCAGGCCAGGCA 8495
QY 1436 GTCATGAACCTACTCTGAGTTATTTACTACTGATTTTGGGGCAGCCTCAGCTATCGTT 1495
DB 8496 GTCATGAACCTACTCTGAGTTATTTACTACTGATTTTGGGGCAGCCTCAGCTATCGTT 8555
QY 1496 TCTTCCACACCTGCTTATGAGAGTATCCATATTTATGCTGCGAGCCAGTAATGCTCCCCA 1555
DB 8556 TCTTCCACACCTGCTTATGAGAGTATCCATATTTATGCTGCGAGC-CAGTAATGCTCCCCA 8614
QY 1556 CGAGATCAGTTTCTGAACTAACCTGGAATTTTATGGGTTTTTATATGCTCCCACTATTA 1615
DB 8615 CGAGATCAGTTTCTGAACTAACCTGGAATTTTATGGGTTTTTATATGCTCCCACTATTA 8674
QY 1616 AATCAACATTTACAGTTCTTCCCTCTGATTTCTCTGTAATAAATAGGCTGCAAAAAA 1675
DB 8675 AATCAACATTTACAGTTCTTCCCTCTGATTTCTCTGTAATAAATAGGCTGCAAAAAA 8734
QY 1676 AAAAAATCTTTTTTAAAAATTAATGCGCATAAAGTATTTGCTCTGGGCTTACTGTATGCTTC 1735
DB 8735 AAAAAATCTTTTTTAAAAATTAATGCGCATAAAGTATTTGCTCTGGGCTTACTGTATGCTTC 8794
QY 1736 TTTTCTTTTCTCTCTTTTCACTAAGTCACTGATTTTATTTAAGATGCGCACTAAT 1795
DB 8795 TTTTCTTTTCTCTCTTTTCACTAAGTCACTGATTTTATTTAAGATGCGCACTAAT 8854
QY 1796 TCAAAACCTATGCTGAGTTCTCAAGGCGAGGTTCACATAGTATGATGAGTTGGGATGGG 1855
DB 8855 TCAAAACCTATGCTGAGTTCTCAAGGCGAGGTTCACATAGTATGATGAGTTGGGATGGG 8914
QY 1856 CTACGGAGAGAAACAGAACTCTAGTTTATTTTAAACCTGATTTTACTGCCACTCC 1915
DB 8915 CTACGGAGAGAAACAGAACTCTAGTTTATTTTAAACCTGATTTTACTGCCACTCC 8974
QY 1916 CCTTAGACTTTGACCATATGACCCCTCGCT-CCCATTTAAGCATAGGGGAGGCTTTAT 1974
DB 8975 CCTTAGACTTTGACCATATGACCCCTCGCT-CCCATTTAAGCATAGGGGAGGCTTTAT 9034

QY	1975	TTTACAAATCGTAATA---GATATCACTTGAGGTGTTTATCAAGAGTTGCGCGGTGGTG	2031
Db	9035	TTTACAAATGGTATAGATGATACACTTGAGGTTTTATCAAGAGTTGCGCGGTGGTG	9094
QY	2032	AAAGTTCAACAACAGAGATTCAGGTTTTGTTGTGCCAGATTCTAAATTTTACATGTTTCTTTT	2091
Db	9095	AAAGTTCAACAACAGAGATTCAGGTTTTGTTGTGCCAGATTCTAAATTTTACATGTTTCTTT	9154
QY	2092	TGCCAAAGGGTGATTTTTTAAATAACATTTGTTTTCTCTTATCTTGTCTTTATTAGGTC	2151
Db	9155	TGCCAAAGGGTGATTTTTTAAATAACATTTGTTTTCTCTTATCTTGTCTTTATTAGGTC	9214
QY	2152	GGAGACCATGAGAAACAGCGTCAAAATCATCTTTTCATGATCCCAAGCTGAAGGCAAGCC	2211
Db	9215	GGAGACCATGAGAAACAGCGTCAAAATCATCTTTTCATGATCCCAAGCTGAAGGCAAGCC	9274
QY	2212	CTCCAGAGACGGTTATGTGTACCCACAAACGAGACACATTTGGTGTACAGACCTTTCGGGGCGTG	2271
Db	9275	CTCCAGAGACGGTTATGTGTACCCACAAACGAGACACATTTGGTGTACAGACCTTTCGGGGCGTG	9334
QY	2272	TCTGAAGCCATAGCCTCCACGGAGAGCCCT	2301
Db	9335	TCTGAAGCCATAGCCTCCACGGAGAGCCCT	9364
RESULT 4			
AAF21289			
ID	AAF21289 standard; DNA; 12461 BP.		
XX			
XX	AAF21289;		
XX			
DT	14-MAR-2001 (first entry)		

CC	(1) can have respiratory, bronchodilator, antiinflammatory, analgesic,
CC	immunosuppressive, antiaesthetic, hypotensive and cytostatic activities.
CC	The antisense oligonucleotides and (1) can be used to down-regulate the
CC	expression and/or activity of target polypeptides associated with
CC	lung/respiratory disorders and malignancies, such as stimulating and
CC	activating peptide factors and transmitters, transcription factors,
CC	immunoglobulins and antibodies, antibody receptors, cytokines and
CC	chemokines, endogenously produced specific and non-specific enzymes,
CC	binding proteins, adhesion molecules and their receptors, cytokine and
CC	chemokine receptors, adenosine receptors, bradykinin receptors, central
CC	nervous system (CNS) and peripheral nervous and non-nervous system
CC	receptors, CNS and peripheral nervous and non-nervous system peptide
CC	transmitters, defensins, growth factors, vasoactive peptides and
CC	receptors, binding proteins and malignancy associated proteins. The
CC	antisense oligonucleotides may be used in this way to treat disorders
CC	including respiratory obstruction (especially pulmonary obstruction
CC	and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or
CC	surfactant hypoproduction which are associated with a disease or
CC	condition selected from pulmonary vasoconstriction, inflammation,
CC	allergies, asthma, impeded respiration, respiratory distress syndrome
CC	(RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC	hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC	pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC	and/or cancer. AAF18434 to AAF1543 represent human polynucleotide
CC	fragments and antisense oligonucleotides used in the exemplification of
CC	the present invention
XX	
SQ	Sequence 12461 BP; 3372 A; 2728 C; 2816 G; 3544 T; 0 U; 1 Other;
	Query Match 97.4%; Score 2241.6; DB 3; Length 12461;
	Best Local Similarity 99.3%; Pred. No. 0;
	Matches 2294; Conservative 1; Mismatches 5; Indels 10; Gaps 4;
QY	1 CTGAATGATGCTCCCAAGTCTATGTGATGAGCTCTTGTGTGCCAGCTGAATAGGTG 60
DB	7056 CTGAATGATGCTCCCAAGTCTATGTGATGAGCTCTTGTGTGCCAGCTGAATAGGTG 7115
QY	61 TGTCCAATGTCATTTTAAAGACTATTATTACACTAATATAGTTTCTTCTCTTTGG 120
DB	7116 TGTCCAATGTCATTTTAAAGACTATTATTACACTAATATAGTTTCTTCTCTCTTTGG 7175
QY	121 ATAATAGGCACTGTGTTCCGTATGGACTTTGGAAGCCCTAGTCCAAGAGAGCCTTGGAGA 180
DB	7176 ATAATAGGCACTGTGTTCCGTATGGACTTTGGAAGCCCTAGTCCAAGAGAGCCTTGGAGA 7235
QY	181 ATTTACTTCCCAAGGCAAGCAGACCGTGAAATATAGATGCCAATGTGCTAGCCAAAAG 240
DB	7236 ATTTACTTCCCAAGGCAAGCAGACCGTGAGAAATAGATGCCAATGTGCTAGCCAAAAG 7295
QY	241 ACAAGAAGTGTGGAAATTTTTCGCAAGCAGGAGAAAGAACTCAGGTGAGCAGAAACACCTT 300
DB	7296 ACAAGAAGTGTGGAAATTTTTCGCAAGCAGGAGAAAGAACTCAGGTGAGCAGAAACACCTT 7355
QY	301 TGTCTTTCAATCAGTTTAAACGCTCTGAACTCCTTCCATATCATGTGTACTGCTCTCTG 360
DB	7356 TGTCTTTCAATCAGTTTAAACGCTCTGAACTCCTTCCATATCATGTGTACTGCTCTCTG 7415
QY	361 TTTTATAGAGACTAAACAGAGACATTTGAAAGTCAGGGTAAGCTGAATATAACATTGCTGA 420
DB	7416 TTTTATAGAGACTAAACAGAGACATTTGAAAGTCAGGGTAAGCTGAATATAACATTGCTGA 7475
QY	421 AATGTTTTTCTGTGTATTTTAAACAGGCTGAAGACTTATGGAGAAAGACTGGAATAA 480
DB	7476 AATGTTTTTCTGTGTATTTTAAACAGGCTCAAGACATTTATGGAGAAAGACTGGAATAA 7535
QY	481 TCATAAGAAAGCAAAAGACTGTTTCAAGCTTCGGAAAAAGTCTATTATTACACAGTTAGT 540
DB	7536 TCATAAGAAAGCAAAAGACTGTTTCAAGCTTCGGAAAAAGTCTATTATTATCACAGTTAGT 7595
QY	541 GAGAGGAAGAAAAATCAGAAGAAAGTTCAAGGAAACCTTAAGACAAAACAGGTAAAGGG 600
DB	7596 GAGAGGAAGAAAAATCAGAAGAAAGTTCAAGGAAACCTTAAGACAAAACAGGTAAAGGG 7655

[illegible]

Db	8735	AAAAATCTTTTAAATAATTTGCCATTAAGTATTTGCTTGGCCCTACTGATGCTTC	8799
Qy	1736	TTTTCTTTTCTCTCTTTTCAACTAAGTCACCGTCAATTTATTAAGATGGCCATAACTAT	1795
Db	8795	TTTTTATTTTCTCTCTTTTCAACTAAGTCACCGTCAATTTATTAAGATGGCCATAACTAT	8854
Qy	1796	TCAAAACCTATCTGAGTTCTCTCAAGGACGGTGCATAGTGTATGAAGCTGGGATGGG	1855
Db	8855	TCAAAACCTATCTGAGTTCTCTCAAGGACGGTGCATAGTGTATGAAGCTGGGATGGG	8914
Qy	1856	CTACGGAAGAACCCAGAACAACTCTAGTTTATTTAAACCTGTATTTACTTGCCCACTTC	1915
Db	8915	CTACGGAAGAACCCAGAACAACTCTAGTTTATTTAAACCTGTATTTACTTGCCCACTTC	8974
Qy	1916	CTTTAGACTTGACCATATGACCCCTCGCTCCCATTTCTAAGCATAGGGGACGCTTTATT	1974
Db	8975	CTTTAGACTTGACCATATGACCCCTCGCTCCCATTTCTAAGCATAGGGGACGCTTTATT	9034
Qy	1975	TTTACAATGGTAATA---GATATCACTTCGAGGTTTTTATCAAAGAGTTGGCGGGTGGTG	2031
Db	9035	TTTACAATGGTAATAGATGATATCACTTCGAGGTTTTTATCAAAGAGTTGGCGGGTGGTG	9094
Qy	2032	AAAGTTTCACACACAGATTCAGGTTTTGTTGTGCCAGATCTTAATTTTACATGTTTCTTT	2091
Db	9095	AAAGTTTCACACACAGATTCAGGTTTTGTTGTGCCAGATCTTAATTTTACATGTTTCTTT	9154
Qy	2092	TGCCAAAGGTTGATTTTTTAAATAAACATTTGTTTTCTCTTATCTTGCTTTATTAGGTC	2151
Db	9155	TGCCAAAGGTTGATTTTTTAAATAAACATTTGTTTTCTCTTATCTTGCTTTATTAGGTC	9214
Qy	2152	GGAGACCATGAGAAAAACAGGTCAAATCATCTTTTTCATGATCCCAAGCTGAAGGCAAGCC	2211
Db	9215	GGAGACCATGAGAAAAACAGGTCAAATCATCTTTTTCATGATCCCAAGCTGAAGGCAAGCC	9274
Qy	2212	CTCCAGAGAGCGTTATGTGACCCACACACCGAGCACATTTGGTGTACAGACCTTCGGGCGCTG	2271
Db	9275	CTCCAGAGAGCGTTATGTGACCCACACACCGAGCACATTTGGTGTACAGACCTTCGGGCGCTG	9334
Qy	2272	TCTGAAGCCATAGCCTCCACGGAGAGCCCT	2301
Db	9335	TCTGAAGCCATAGCCTCCACGGAGAGCCCT	9364
RESULT 5			
ABK83778			
ID	ABK83778 standard; cdna; 12461 BP.		
XX	ABK83778;		
AC	ABK83778;		
XX	14-AUG-2002 (first entry)		
DT	Human cdna differentially expressed in granulocytic cells #349.		
XX	Human; ss; granulocytic cell; DNA chip; bacterial infection;		
XX	viral infection; parasitic infection; protozoal infection;		
KW	fungal infection; sterile inflammatory disease; psoriasis;		
KW	rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;		
KW	cardiac reperfusion injury; renal reperfusion injury; AKDS;		
KW	adult respiratory distress syndrome; inflammatory bowel disease;		
KW	Crohn's disease; ulcerative colitis; periodontal disease;		
XX	granulocyte activation; chronic inflammation; allergy.		
OS	Homo sapiens.		
XX	WO200228999-A2.		
PN	11-APR-2002.		
XX	03-OCT-2001; 2001WO-US030821.		
PF	03-OCT-2000; 2000US-0237189P.		
XX			
PR			
XX			

10

(GENE-) GENE LOGIC INC.

Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;

WPI; 2002-435328/46.

Detecting granulocyte activation by detecting differential expression of genes associated with granulocyte activation, which serves as diagnostic markers that is useful for monitoring disease states and drug toxicity.

Claim 1; SEQ ID NO 349; 114pp; English.

The invention relates to detecting (M1) granulocyte (GC) activation (GCA), by detecting the level of expression of gene(s) identified by DNA chip analysis as given in the specification, and comparing the expression level to an expression level in an unactivated GC, where differential expression of Gs is indicative of GCA. Also included are modulating (M2) Gs by contacting GC with an agent that alters the expression of at least one gene in Gs; (2) screening (M3) for an agent capable of modulating GCA or an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease using the gene expression profile; (3) detecting (M4) an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by detecting the level of expression in a sample of the tissue of gene(s) from Gs, where the level of expression of the gene is indicative of inflammation; (4) treating (M5) an inflammation (especially chronic) or in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by contacting a tissue having inflammation with an agent that modulates the expression of gene(s) from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for modulating GCA; M3 is useful for screening an agent capable of modulating GCA preferably in an inflammation in a tissue; M4 is useful for detecting an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease (e.g. psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal reperfusion injury, ARDS, adult respiratory distress syndrome, inflammatory bowel disease, Crohn's disease, ulcerative colitis, periodontal disease; also bacterial infection, viral infection, parasitic infection, protozoal infection, fungal infection and M5 is useful for treating one of the above conditions. The present sequence represents a gene differentially expressed in granulocytes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 12461 BP; 3372 A; 2728 C; 2816 G; 3544 T; 0 U; 1 Other;

Query Match 97.4%; Score 2241.6; DB 6; Length 12461;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2294; Conservative 1; Mismatches 5; Indels 10; Gaps 4;

QY	1	CTGAAATGATGCTCCCAAGTCTATGATGAGTCTCTTGTCGCCAGTGGAAATAGGTG	60
DB	7056	CTGAAATGATGCTCCCAAGTCTATGATGAGTCTCTTGTCGCCAGTGGAAATAGGTG	7115
QY	61	TGTCATGTCATTTTAAAGCTATTAAATACACTAATATAGTTCTTCTCTCTTTGG	120
DB	7116	TGTCATGTCATTTTAAAGCTATTAAATACACTAATATAGTTCTTCTCTCTTTGG	7175
QY	121	ATAATAGACAGTGTTCCTGATGAGTCTGGAAGCCCTAGGTCAAGAGAGCTTTGAGA	180
DB	7176	ATAATAGACAGTGTTCCTGATGAGTCTGGAAGCCCTAGGTCAAGAGAGCTTTGAGA	7235
QY	181	ATTTACTTCCCAAGGACACACACGCTGAAATAGATGCTAGTGGTCCCAAAAG	240
DB	7236	ATTTACTTCCCAAGGACACACGCTGAAATAGATGCTAGTGGTCCCAAAAG	7295
QY	241	ACAAGAAGTGTGGAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAGAACCTT	300
DB	7296	ACAAGAAGTGTGGAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAGAACCTT	7355

QY	301	TGCTTTTCAATCAGTTTAAACAGCCTCCTGAACTCCTTCTCTATCATGGTACTGCCTTCTTG	360
DB	7356	TGCTTTTCAATCAGTTTAAACAGCCTCCTGAACTCCTTCTCTATCATGGTACTGCCTTCTTG	7415
QY	361	TTTATAGAGAGCTAAACAGAGACATTGAAAGTCAAGGTAAAGCTGAATATAACATTGCTGA	420
DB	7416	TTTATAGAGAGCTAAACAGAGACATTGAAAGTCAAGGTAAAGCTGAATATAACATTGCTGA	7475
QY	421	AATGTTTTTCTTGCTGTTTAAACAGGCTCAAGACATTTATGGAGAAAGCTGGAATAA	480
DB	7476	AATGTTTTTCTTGCTGTTTAAACAGGCTCAAGACATTTATGGAGAAAGCTGGAATAA	7535
QY	481	TCATAGAAAGGAAAGACCTGTTTCAAGCTTGGGAAAAAGTGTATTTATCAGCAGTTAGT	540
DB	7536	TCATAGAAAGGAAAGACCTGTTTCAAGCTTGGGAAAAAGTGTATTTATCAGCAGTTAGT	7595
QY	541	GAGAGGAAAGAAATCAGAAAGTTCAGAGGAAACACCTAAGACAAACAGGCTAAGAGGG	600
DB	7596	GAGAGGAAAGAAATCAGAAAGTTCAGAGGAAACACCTAAGACAAACAGGCTAAGAGGG	7655
QY	601	AAGGAGGAAAGAAATAGGTAAGAGGTTTCAAGAAACAACTAGTCCCGCAGTCAGTGATGCCAG	660
DB	7656	AAGGAGGAAAGAAATAGGTAAGAGGTTTCAAGAAACAACTAGTCCCGCAGTCAGTGATGCCAG	7715
QY	661	CAGCTGTTTCTTCCAGCCCTTCTTACCGGCGAGGTGAAGACTTAGAAAAACAGTAGCAG	720
DB	7716	CAGCTGTTTCTTCCAGCCCTTCTTACCGGCGAGGTGAAGACTTAGAAAAACAGTAGCAG	7775
QY	721	AGGAGATCTATCATCTTATAGATTTAAAGAGGAGGAAAGAACTCCCTCTTAAATATTCCA	780
DB	7776	AGGAGATCTATCATCTTATAGATTTAAAGAGGAGGAAAGAACTCCCTCTTAAATATTCCA	7835
QY	781	TGAAGCTCTGGAATCAAAACCGATCTCTCTGTATTTTAGCACATACATTTCACTTAC	840
DB	7836	TGAAGCTCTGGAATCAAAACCGATCTCTCTGTATTTTAGCACATACATTTCACTTAC	7895
QY	841	AGGTAGATTTTCCCAACCAAAATATATCCAGAGATGCTCTTGTTCATTTGGTGTATATACAGC	900
DB	7896	AGGTAGATTTTCCCAACCAAAATATATCCAGAGATGCTCTTGTTCATTTGGTGTATATACAGC	7955
QY	901	CTTTGCCCTCTGAGTCAATGATTTTATTTACCTTTCCCTGAGAAATCGAAAAATCAATTTGG	960
DB	7956	CTTTGCCCTCTGAGTCAATGATTTTATTTACCTTTCCCTGAGAAATCGAAAAATCAATTTGG	8015
QY	961	GGAGCGGACATTTAGAAAAGAAATCAAAAGTGTCTAGGATATCAAAATCTTCAATAGTT	1020
DB	8016	GGAGCGGACATTTAGAAAAGAAATCAAAAGTGTCTAGGATATCAAAATCTTCAATAGTT	8075
QY	1021	GCAGTTTATTCAGATGCGCCAAAGGAAAAATAAAGTCAATAGATAGGTTGGTAGAATTTAG	1080
DB	8076	GCAGTTTATTCAGATGCGCCAAAGGAAAAATAAAGTCAATAGATAGGTTGGTAGAATTTAG	8135
QY	1081	AACATGCTGTTTTTTCAGGTTTTATGTC-----TTTTTTTTTTTTTTTTTTTTTTTTTAAATAGGG	1135
DB	8136	AACATGCTGTTTTTTCAGGTTTTATGTC-----TTTTTTTTTTTTTTTTTTTTTTTTTAAATAGGG	8195
QY	1136	AAATGTTTTTGGTGAGAGCCCAATGTCATTTCCAAAAGCTCTCTTTTCTCTGTCAGTC	1195
DB	8196	AAATGTTTTTGGTGAGAGCCCAATGTCATTTCCAAAAGCTCTCTTTTCTCTGTCAGTC	8255
QY	1196	ATGTCCTGGGACAGAAAGGATCTGGATTAGGCAACATCATAGAGTTGCTCTGAGCTGC	1255
DB	8256	ATGTCCTGGGACAGAAAGGATCTGGATTAGGCAACATCATAGAGTTGCTCTGAGCTGC	8315
QY	1256	TCCTGTTGGTGAATAACCTTCCAAATCCTAAACCTTTTGGAAATTCACAGCTCAAAAGGAGA	1315
DB	8316	TCCTGTTGGTGAATAACCTTCCAAATCCTAAACCTTTTGGAAATTCACAGCTCAAAAGGAGA	8375
QY	1316	AACCTACTCTCTGATCTACCATGTTCTGCAATTTTCTCATGGTCTATGGAACCTTC	1375
DB	8376	AACCTACTCTCTGATCTACCATGTTCTGCAATTTTCTCATGGTCTATGGAACCTTC	8435

QY	1376	TCCTAGAAATCCAGTGGCAAGATTCCTATGATTAAAGTGTCTGAGCTCAGGCCAGGCA	1435	DE	Human nucleic acid sequence.
Db	8436	TCCTAGAAATCCAGTGGCAAGATTCCTATGATTAAAGTGTCTGAGCTCAGGCCAGGCA	8495	XX	Human; antisense; lung dysfunction; nasal airway dysfunction;
QY	1436	GTCAATCACTACTCTGAGTATTTACTACTGATTTGTGGGCGACCTCAGCTATCGTT	1495	KW	antiinflammatory steroid; ubiquinone; antinflammatory; antiallergic;
Db	8496	GTCAATCACTACTCTGAGTATTTACTACTGATTTGTGGGCGACCTCAGCTATCGTT	8555	KW	antiaesthetic; hypotensive; immunosuppressive; cytostatic; gene therapy;
QY	1496	TCCTACACCTCTTATGAGATATCCATATTTATGGTCGAGGCGCAATATGCTCCCA	1555	KW	antisense gene therapy; respiratory; lung; adenosine sensitivity;
Db	8556	TCCTACACCTCTTATGAGATATCCATATTTATGGTCGAGGCGCAATATGCTCCCA	8614	KW	adenosine receptor; bronchodilation; bronchoconstriction; lung allergy;
QY	1556	CGAGATCAGTTTCTGAACCTAAGTATTTTATGAGGTTTATTAATGCACTATTA	1615	XX	lung inflammation; respiratory disease; ds.
Db	8615	CGAGATCAGTTTCTGAACCTAAGTATTTTATGAGGTTTATTAATGCACTATTA	8674	OS	Homo sapiens.
QY	1616	AATCAACATTTACAGTTCTTCCCTCTGTTTCTCTGTAACATTTAGGCTGCAAAAA	1675	XX	WO200285308-A2.
Db	8675	AATCAACATTTACAGTTCTTCCCTCTGTTTCTCTGTAACATTTAGGCTGCAAAAA	8734	PN	31-OCT-2002.
QY	1676	AAAAATCTTTTAAAAATTAATGCAATAAGTATTTGCTGGGCCCTACTGTATGCTTC	1735	PD	23-APR-2002; 2002WO-US013135.
Db	8735	AAAAATCTTTTAAAAATTAATGCAATAAGTATTTGCTGGGCCCTACTGTATGCTTC	8794	PF	24-APR-2001; 2001US-0286137P.
QY	1736	TTTTCTTTTCTCTCTTTCACTAAGTCACCGTCAATTTATTAAGATGGCCAACTAT	1795	PR	(EPIC-) EPIGENESIS PHARM INC.
Db	8795	TTTTCTTTTCTCTCTTTCACTAAGTCACCGTCAATTTATTAAGATGGCCAACTAT	8854	PA	Nyce JW, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D;
QY	1796	TCAAAACCTATGCTGAGTTCCTCAAGCGAGGTGCATAGTATGATGAAGTTGGATGGG	1855	PI	Miller S, Tang L, Shahabuddin S;
Db	8855	TCAAAACCTATGCTGAGTTCCTCAAGCGAGGTGCATAGTATGATGAAGTTGGATGGG	8914	XX	WPI; 2003-229219/22.
QY	1856	CTACGGAAGAACCGAACAACCTAGTTTATTTAAACCTGTATTTACTGCCACTCC	1915	XX	Pharmaceutical composition for treating ailments associated with impaired
Db	8915	CTACGGAAGAACCGAACAACCTAGTTTATTTAAACCTGTATTTACTGCCACTCC	8974	XX	respiration, has oligo(s) antisense to specific gene(s) or its
QY	1916	CCTTAGACTTGACCATATGACCCCTCGCT - CCCATTCTAAGCATAGGCGCAGGCTTTATT	1974	PT	corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or
Db	8975	CCTTAGACTTGACCATATGACCCCTCGCT - CCCATTCTAAGCATAGGCGCAGGCTTTATT	9034	PT	ubiquinone.
QY	1975	TTTCAATGGTAAATA - -GATATCACTTGAGGTTTATCAAGATTTGGCGGGTGGTG	2031	PS	Disclosure; SEQ ID NO 12225; 872pp; English.
Db	9035	TTTCAATGGTAAATAGATATATCACTTGAGGTTTATCAAGATTTGGCGGGTGGTG	9094	XX	The invention relates to a novel pharmaceutical composition, which has a
QY	2032	AAAGTTTCAACACGATTCAGGTTTGTGTGCGAGATTCATATTTTACATGTTCTTT	2091	CC	first active agent comprising an oligonucleotide antisense to the
Db	9095	AAAGTTTCAACACGATTCAGGTTTGTGTGCGAGATTCATATTTTACATGTTCTTT	9154	CC	initiation codon, coding region, 5' or 3' end genomic flanking regions,
QY	2092	TGCCAAAGGGTGATTTTAAATAACAATTTGTTTCTCTATCTTCTTTATAGTGC	2151	CC	5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of
Db	9155	TGCCAAAGGGTGATTTTAAATAACAATTTGTTTCTCTATCTTCTTTATAGTGC	9214	CC	junctions of genes encoding a polypeptide associated with lung and/or
QY	2152	GGAGACCATGAGAACAGCGTCAATCATCTTTTCATGATCCCAAGCTGAAGGCAAGCC	2211	CC	nasal airway dysfunction and a second active agent comprising an
Db	9215	GGAGACCATGAGAACAGCGTCAATCATCTTTTCATGATCCCAAGCTGAAGGCAAGCC	9274	CC	antiinflammatory steroid and ubiquinone. A composition of the invention
QY	2212	CTCCAGAGAGCGTTATGTGACCCCAACACCGACACATTTGTTGACAGACCTTCGGGCGCTG	2271	CC	has antiinflammatory, antiallergic, antiasthmatic, hypotensive, or
Db	9275	CTCCAGAGAGCGTTATGTGACCCCAACACCGACACATTTGTTGACAGACCTTCGGGCGCTG	9334	CC	immunosuppressive, and cytostatic activity. The composition may have a
QY	2272	TCTGAAGCCATAGCTCCACGAGAGCCCT	2301	CC	use in antisense gene therapy. The composition is useful for treating or
Db	9335	TCTGAAGCCATAGCTCCACGAGAGCCCT	9364	CC	preventing a respiratory, lung or malignant disease or condition, also
QY	2301	CTGAATGATGCTCCCAAGTGCATGATGAGTCTCTTGTGCGCCAGTGAATAGGTG	7115	CC	for enhancing the prophylactic or therapeutic respiratory effect of an
Db	7115	CTGAATGATGCTCCCAAGTGCATGATGAGTCTCTTGTGCGCCAGTGAATAGGTG	7175	CC	antiinflammatory steroid in a subject, for reducing or depleting levels
QY	7175	TGTCATGTCATTTTAAAGACTATTAATACACTAATATAGTTCTCTCTCTTGG	7235	CC	of, or reducing sensitivity to adenosine, reducing levels of adenosine
Db	7235	TGTCATGTCATTTTAAAGACTATTAATACACTAATATAGTTCTCTCTCTTGG	7295	CC	receptor, producing bronchodilation, increasing levels of ubiquinone or
QY	7295	ATATAGGACAGTGTTCGGTATGGACTTGGAGCCCTAGTCCAGAGAGCCCTTGAGA	7355	CC	lung surfactant in a subject's tissue, or treating bronchoconstriction,
Db	7355	ATATAGGACAGTGTTCGGTATGGACTTGGAGCCCTAGTCCAGAGAGCCCTTGAGA	7415	CC	lung inflammation, lung allergies, or a respiratory disease or condition.
QY	7415	ATATAGGACAGTGTTCGGTATGGACTTGGAGCCCTAGTCCAGAGAGCCCTTGAGA	7475	CC	Note: The sequence data for this patent is not represented in the printed
Db	7475	ATATAGGACAGTGTTCGGTATGGACTTGGAGCCCTAGTCCAGAGAGCCCTTGAGA	7535	CC	specification, but was obtained in electronic format directly from WIPO
QY	7535	ATTTACTTCCCAAAAGGCAACAGCCGTGCAAAATAGATCCCAATGTGCTAGCAAAAAG	7595	CC	at ftp.wipo.int/pub/published_pct_sequences
Db	7595	ATTTACTTCCCAAAAGGCAACAGCCGTGCAAAATAGATCCCAATGTGCTAGCAAAAAG	7655	XX	Sequence 12461 BP; 3372 A; 2728 C; 2816 G; 3544 T; 0 U; 1 Other;
QY	7655	CTGAATGATGCTCCCAAGTGCATGATGAGTCTCTTGTGCGCCAGTGAATAGGTG	7715	XX	Query Match 97.4%; Score 2241.6; DB 7; Length 12461;
Db	7715	CTGAATGATGCTCCCAAGTGCATGATGAGTCTCTTGTGCGCCAGTGAATAGGTG	7775	XX	Best Local Similarity 99.3%; Pred. No. 0;
QY	7775	TGTCATGTCATTTTAAAGACTATTAATACACTAATATAGTTCTCTCTCTTGG	7835	XX	Matches 2294; Conservative 1; Mismatches 5; Indels 10; Gaps 4;
Db	7835	TGTCATGTCATTTTAAAGACTATTAATACACTAATATAGTTCTCTCTCTTGG	7895	XX	
QY	7895	ATATAGGACAGTGTTCGGTATGGACTTGGAGCCCTAGTCCAGAGAGCCCTTGAGA	7955	XX	
Db	7955	ATATAGGACAGTGTTCGGTATGGACTTGGAGCCCTAGTCCAGAGAGCCCTTGAGA	8015	XX	
QY	8015	ATTTACTTCCCAAAAGGCAACAGCCGTGCAAAATAGATCCCAATGTGCTAGCAAAAAG	8075	XX	
Db	8075	ATTTACTTCCCAAAAGGCAACAGCCGTGCAAAATAGATCCCAATGTGCTAGCAAAAAG	8135	XX	
QY	8135	CTGAATGATGCTCCCAAGTGCATGATGAGTCTCTTGTGCGCCAGTGAATAGGTG	8195	XX	
Db	8195	CTGAATGATGCTCCCAAGTGCATGATGAGTCTCTTGTGCGCCAGTGAATAGGTG	8255	XX	
QY	8255	TGTCATGTCATTTTAAAGACTATTAATACACTAATATAGTTCTCTCTCTTGG	8315	XX	
Db	8315	TGTCATGTCATTTTAAAGACTATTAATACACTAATATAGTTCTCTCTCTTGG	8375	XX	
QY	8375	ATATAGGACAGTGTTCGGTATGGACTTGGAGCCCTAGTCCAGAGAGCCCTTGAGA	8435	XX	
Db	8435	ATATAGGACAGTGTTCGGTATGGACTTGGAGCCCTAGTCCAGAGAGCCCTTGAGA	8495	XX	
QY	8495	ATTTACTTCCCAAAAGGCAACAGCCGTGCAAAATAGATCCCAATGTGCTAGCAAAAAG	8555	XX	
Db	8555	ATTTACTTCCCAAAAGGCAACAGCCGTGCAAAATAGATCCCAATGTGCTAGCAAAAAG	8615	XX	
QY	8615	CTGAATGATGCTCCCAAGTGCATGATGAGTCTCTTGTGCGCCAGTGAATAGGTG	8675	XX	
Db	8675	CTGAATGATGCTCCCAAGTGCATGATGAGTCTCTTGTGCGCCAGTGAATAGGTG	8735	XX	
QY	8735	TGTCATGTCATTTTAAAGACTATTAATACACTAATATAGTTCTCTCTCTTGG	8795	XX	
Db	8795	TGTCATGTCATTTTAAAGACTATTAATACACTAATATAGTTCTCTCTCTTGG	8855	XX	
QY	8855	ATATAGGACAGTGTTCGGTATGGACTTGGAGCCCTAGTCCAGAGAGCCCTTGAGA	8915	XX	
Db	8915	ATATAGGACAGTGTTCGGTATGGACTTGGAGCCCTAGTCCAGAGAGCCCTTGAGA	8975	XX	
QY	8975	ATTTACTTCCCAAAAGGCAACAGCCGTGCAAAATAGATCCCAATGTGCTAGCAAAAAG	9035	XX	
Db	9035	ATTTACTTCCCAAAAGGCAACAGCCGTGCAAAATAGATCCCAATGTGCTAGCAAAAAG	9095	XX	
QY	9095	CTGAATGATGCTCCCAAGTGCATGATGAGTCTCTTGTGCGCCAGTGAATAGGTG	9155	XX	
Db	9155	CTGAATGATGCTCCCAAGTGCATGATGAGTCTCTTGTGCGCCAGTGAATAGGTG	9215	XX	
QY	9215	TGTCATGTCATTTTAAAGACTATTAATACACTAATATAGTTCTCTCTCTTGG	9275	XX	
Db	9275	TGTCATGTCATTTTAAAGACTATTAATACACTAATATAGTTCTCTCTCTTGG	9335	XX	
QY	9335	ATATAGGACAGTGTTCGGTATGGACTTGGAGCCCTAGTCCAGAGAGCCCTTGAGA	9395	XX	
Db	9395	ATATAGGACAGTGTTCGGTATGGACTTGGAGCCCTAGTCCAGAGAGCCCTTGAGA	9455	XX	
QY	9455	ATTTACTTCCCAAAAGGCAACAGCCGTGCAAAATAGATCCCAATGTGCTAGCAAAAAG	9515	XX	
Db	9515	ATTTACTTCCCAAAAGGCAACAGCCGTGCAAAATAGATCCCAATGTGCTAGCAAAAAG	9575	XX	
QY	9575	CTGAATGATGCTCCCAAGTGCATGATGAGTCTCTTGTGCGCCAGTGAATAGGTG	9635	XX	
Db	9635	CTGAATGATGCTCCCAAGTGCATGATGAGTCTCTTGTGCGCCAGTGAATAGGTG	9695	XX	
QY	9695	TGTCATGTCATTTTAAAGACTATTAATACACTAATATAGTTCTCTCTCTTGG	9755	XX	
Db	9755	TGTCATGTCATTTTAAAGACTATTAATACACTAATATAGTTCTCTCTCTTGG	9815	XX	
QY	9815	ATATAGGACAGTGTTCGGTATGGACTTGGAGCCCTAGTCCAGAGAGCCCTTGAGA	9875	XX	
Db	9875	ATATAGGACAGTGTTCGGTATGGACTTGGAGCCCTAGTCCAGAGAGCCCTTGAGA	9935	XX	
QY	9935	ATTTACTTCCCAAAAGGCAACAGCCGTGCAAAATAGATCCCAATGTGCTAGCAAAAAG	10000	XX	
Db	10000	ATTTACTTCCCAAAAGGCAACAGCCGTGCAAAATAGATCCCAATGTGCTAGCAAAAAG			

RESULT 6
ID ABZ96983
AC ABZ96983; standard; DNA; 12461 BP.
XX
XX
XX
XX
XX
XX
17-OCT-2003 (first entry)

[illegible]

8316	Db	TCTTTGGTGATAAACCCCTTCCAAATCTCAAACCTTTTTTGGAAATTCACAAGCTCAAAGGAGGA	8375
1316	Qy	AACCTACTCTCTGATCTTACCACATGTTCTGCAATTTTCTATCATGTGCTATGAAACTTC	1375
8376	Db	AACCTACTCTCTGATCTTACCACATGTTCTGCAATTTTCTATCATGTGCTATGAAACTTC	8435
1376	Qy	TCATTAGAAATCCAGTGGCAGAGAGTTCTATGATTAAGTGTCTGAGCTCAGGCCAGGCA	1435
8436	Db	TCATTAGAAATCCAGTGGCAGAGAGTTCTATGATTAAGTGTCTGAGCTCAGGCCAGGCA	8495
1436	Qy	GTCATGAACCTACTTCTCAGTTATTCTACTACTGATTTGTGGGGCAGCCTCAGCTATCGGTT	1495
8496	Db	GTCATGAACCTACTTCTCAGTTATTCTACTACTGATTTGTGGGGCAGCCTCAGCTATCGGTT	8555
1496	Qy	TCATTACACCTGCTTATGAGAGTATCCATATTTATGGTCGCAGGCCAGTAATGCTCCCA	1555
8556	Db	TCATTACACCTGCTTATGAGAGTATCCATATTTATGGTCGCAGG-CAGTAATGCTCCCA	8614
1556	Qy	CGAGATCAGTTTCTGAACCTGAACTTTTATGGGTTTTTATGCGGTTTTTATGCCAATATTA	1615
8615	Db	CGAGATCAGTTTCTGAACCTGAACTTTTATGGGTTTTTATGCGGTTTTTATGCCAATATTA	8674
1616	Qy	AATCAACATTTACAGTTCTTCCCTCTGATTTCTCCCTGTTAAACAATTAGGCTGCAAAAA	1675
8675	Db	AATCAACATTTACAGTTCTTCCCTCTGATTTCTCCCTGTTAAACAATTAGGCTGCAAAAA	8734
1676	Qy	AAAAAATCTTTTTTAAAAATAATTGCCATAAAGATTTGCTCTGGGCTACTGTATGCTTC	1735
8735	Db	AAAAAATCTTTTTTAAAAATAATTGCCATAAAGATTTGCTCTGGGCTACTGTATGCTTC	8794
1736	Qy	TTTTTCTTTTCTCTCTTTCTCACTAGTACCGTCAATTTATAGATGCCATACACTAT	1795
8795	Db	TTTTTCTTTTCTCTCTTTTCAACTAGTACCGTCAATTTATAGATGCCATACACTAT	8854
1796	Qy	TCAAAACCTATGCTGAGTTCTCTCAAGCGAGGGTCAATAGTATGAAGTTGGATGGGG	1855
8855	Db	TCAAAACCTATGCTGAGTTCTCTCAAGCGAGGGTCAATAGTATGAAGTTGGATGGGG	8914
1856	Qy	CTACGGAAGAACCGAGAACACTCTAGTTTATTAAACCTGTATTACTGCCCCCTCC	1915
8915	Db	CTACGGAAGAACCGAGAACACTCTAGTTTATTAAACCTGTATTACTGCCCCCTCC	8974
1916	Qy	CCATTAGACTTTGACCATATGACCCCTCGCT-CCATTCTAAGCATAGGGCAGGCTTTAT	1974
8975	Db	CCATTAGACTTTGACCATATGACCCCTCGCT-CCATTCTAAGCATAGGGCAGGCTTTAT	9034
1975	Qy	TTTACATGCGTATA--GATATCACTCAGGTTTTATCAAGAGTTGCGGGGGTGGT	2031
9035	Db	TTTACATGCGTATAATGATGATATCACTCAGGTTTTATCAAGAGTTGCGGGGGTGGT	9094
2032	Qy	AAAGTTTCACAAACAGATTTCAGGTTTTGTGTGCCAGATTTCAATTTATCATGTTCTTT	2091
9095	Db	AAAGTTTCACAAACAGATTTCAGGTTTTGTGTGCCAGATTTCAATTTATCATGTTCTTT	9154
2092	Qy	TGCCAAGGGTGATTTTTTTTAAATAAACATTTGTTTCTTATCTGCTTTATTAGTCT	2151
9155	Db	TGCCAAGGGTGATTTTTTTTAAATAAACATTTGTTTCTTATCTGCTTTATTAGTCT	9214
2152	Qy	GGAGACCATGAGAAACAGCGTCAAAATCATCTTTTCATGATCCCAAGCTGAAAGCAAGC	2211
9215	Db	GGAGACCATGAGAAACAGCGTCAAAATCATCTTTTCATGATCCCAAGCTGAAAGCAAGC	9274
2212	Qy	CTCCAGAGCGTTATGTACCCACAAACCGACATTTGGTCACAGACCTTCGGGGCTG	2271
9275	Db	CTCCAGAGCGTTATGTACCCACAAACCGACATTTGGTCACAGACCTTCGGGGCTG	9334
2272	Qy	TCTGAAGCCATAGCCTCCACGGAGAGCCCT	2301
9335	Db	TCTGAAGCCATAGCCTCCACGGAGAGCCCT	9364

ACA64929 standard; DNA; 12461 BP.
ACA64929;
27-JUN-2003 (first entry)
Human EDN1 DNA corresponding to J05008.
Human; chronic inflammatory joint disease; infection; tumour;
antiflammarory; cycostatic; antirheumatic; antineumatic; ds.
immunosuppressive; gene therapy; etiological pathogenicity; ds.
Homo sapiens.
DE10127572-A1.
05-DEC-2002.
30-MAY-2001; 2001DE-01027572.
30-MAY-2001; 2001DE-01027572.
(PATH-) PATHARRAY GNBH.
Haeupl T, Ungethuen U, Blaess S;
WPI; 2003-240797/24.
Reagents for diagnosis, study and therapy of chronic inflammatory joint
and other diseases, comprises any of many specified genes or derived
proteins.
Claim 1; Page; 12pp; German.
This invention describes a novel reagent for diagnosis, molecular
definition and therapy of chronic inflammatory joint diseases, and other
inflammatory disorders, infective or tumour diseases in humans. The
products of the invention have antiinflammatory, cytostatic,
antiarthritic, antirheumatic and immunosuppressive activity and can be
used for gene therapy. The reagent of the invention and any proteins and
antibodies derived from it, are used (i) for analysing tissue and blood
samples for medical diagnosis; (ii) for diagnosis and characterisation of
chronic joint diseases, on the basis of molecular characterisation, and
determining the etiological pathogenicity principle of as yet
uncharacterised inflammatory diseases, also monitoring progression and/or
treatment of disease, and optimisation of therapy and (iii) for
developing treatments for inflammatory diseases, particularly of joints,
infections and tumours. ACA64801-ACA64965 represent human polynucleotides
used in the method of the invention
Sequence 12461 BP; 3372 A; 2728 C; 2816 G; 3544 T; 0 U; 1 Other;
Query Match 97.4%; Score 2241.6; DB 7; Length 12461;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2294; Conservative 1; Mismatches 5; Indels 10; Gaps 4;
QY 1 CTGAATGATGCTCCCAAGTGTATCTGATGAGCTCCTGTGTGCTCCAGTGGATAGTG 60
DB 7056 CTGAATGATGCTCCCAAGTGTATCTGATGAGCTCCTGTGTGCTCCAGTGGATAGTG 7115
QY 61 TGTCCATGTGCTATTTAAAGACTATTAATACACTAATATAGTTCCTCTCTTTGG 120
DB 7116 TGTCCATGTGCTATTTAAAGACTATTAATACACTAATATAGTTCCTCTCTTTGG 7175
QY 121 ATAATAGGACCTGTTCCGATGACTTGGAGCCCTAGTCCCAAGAGACCTTGGAGA 180
DB 7176 ATAATAGGACCTGTTCCGATGACTTGGAGCCCTAGTCCCAAGAGACCTTGGAGA 7235
QY 181 ATTTACTTCCCAAGGCAACAGACCGTGAATAATAGATGCCAATGTCTAGCCAAAAG 240
DB 7236 ATTTACTTCCCAAGGCAACAGACCGTGAATAATAGATGCCAATGTCTAGCCAAAAG 7295
QY 241 ACAAGAAGTGTGGAAATTTTCCCAAGGCAAGAAAGAACTCAGGTGACGAAACACCTT 300

DB 7296 ACAGAAGTGTGGAAATTTTCCCAAGGCAAGAAAGAACTCAGGTGACGAAACACCTT 7355
QY 301 TGTCTTTCAATCAGTTTAAAGCTTCTGAACTTCTTCTCTCTATCATGTACTGCTTCTG 360
DB 7356 TGTCTTTCAATCAGTTTAAAGCTTCTGAACTTCTTCTCTATCATGTACTGCTTCTG 7415
QY 361 TTTTAGAGAGACTAACAGAGACATTGAAAGTCAAGGTAAAGCTGAATAATAAATCTCTGA 420
DB 7416 TTTTAGAGAGACTAACAGAGACATTGAAAGTCAAGGTAAAGCTGAATAATAAATCTCTGA 7475
QY 421 AATGTTTTCTTCTGTTGTTATTTTAAAGGCTGAAGACATTATGGAAGAAAGACTGGAATAA 480
DB 7476 AATGTTTTCTTCTGTTGTTATTTTAAAGGCTGAAGACATTATGGAAGAAAGACTGGAATAA 7535
QY 481 TCATAAGAAAGGAAGAGACTGTTTCCCAAGCTTGGGAAAGAGTGTATTTATCAGCAGTAGT 540
DB 7536 TCATAAGAAAGGAAGAGACTGTTTCCCAAGCTTGGGAAAGAGTGTATTTATCAGCAGTAGT 7595
QY 541 GAGAGGAAGAAAGAAATCAGAAGAGTTTCAAGGAACACCTTAAGACAAACCCAGGTAAAGGG 600
DB 7596 GAGAGGAAGAAAGAAATCAGAAGAGTTTCAAGGAACACCTTAAGACAAACCCAGGTAAAGGG 7655
QY 601 AAGGAAGAAAGAAATAGGTAAGAGTTTCAAGGAACACCTTAAGACAAACCCAGGTAAAGGG 660
DB 7656 AAGGAAGAAAGAAATAGGTAAGAGTTTCAAGGAACACCTTAAGACAAACCCAGGTAAAGGG 7715
QY 661 CAGCTGTTTCTCCTCAGCCCTTCTTACCCGGGAGGTGAAAGACTTAGAAAAACAGTAGCAG 720
DB 7716 CAGCTGTTTCTCCTCAGCCCTTCTTACCCGGGAGGTGAAAGACTTAGAAAAACAGTAGCAG 7775
QY 721 AGGAGATCTATGTCATCTATAGATTAAAGAGGACAAAGAAATCCCTCTTAAATATTCCA 780
DB 7776 AGGAGATCTATGTCATCTATAGATTAAAGAGGACAAAGAAATCCCTCTTAAATATTCCA 7835
QY 781 TGAAGCTCTGGAATGAAACCGGATGCTCTGTACCTTTTAGACATACCATCTTCACTAC 840
DB 7836 TGAAGCTCTGGAATGAAACCGGATGCTCTGTACCTTTTAGACATACCATCTTCACTAC 7895
QY 841 AGGTAGATTTCCCAACCAAAATATATCCAGAGATGCTTTGTCAATGGGTATATACAGC 900
DB 7896 AGGTAGATTTCCCAACCAAAATATATCCAGAGATGCTTTGTCAATGGGTATATACAGC 7955
QY 901 CTTTGGCTCTCTCAGTCAATGTTTACACATTTCCCTCAGAAATCGAAATCATTTTGG 960
DB 7956 CTTTGGCTCTCTCAGTCAATGTTTACACATTTCCCTCAGAAATCGAAATCATTTTGG 8015
QY 961 GGAGCGGACATTTAGAAAAAGAAATCAAAGTGTGATGATAATCAAATTTCTTCAATAAGTT 1020
DB 8016 GGAGCGGACATTTAGAAAAAGAAATCAAAGTGTGATGATAATCAAATTTCTTCAATAAGTT 8075
QY 1021 GCAGTTATTCAGATGGCCAAAGAAAAATAAAGTCAATAGATGGGTTGGTAGAATTTAG 1080
DB 8076 GCAGTTATTCAGATGGCCAAAGAAAAATAAAGTCAATAGATGGGTTGGTAGAATTTAG 8135
QY 1081 AACATGCTGTTTTTCAAGTTTATGCTC-----TTTTTTTTTTTTTTTTTTTTTAAATAGGG 1135
DB 8136 AACATGCTGTTTTTCAAGTTTATGCTC-----TTTTTTTTTTTTTTTTTTTTTAAATAGGG 8195
QY 1136 AATGTTTGGTGGAGAGCAATGTCATTTCAAAGAGTCTCTCTTTTCTCGGTGAGTC 1195
DB 8196 AATGTTTGGTGGAGAGCAATGTCATTTCAAAGAGTCTCTCTTTTCTCGGTGAGTC 8255
QY 1196 ATGTGCTGGGACAGAGAGAGGATCTGGATTAGGCAACATCATAGAGTTGCTCTGAGCTGC 1255
DB 8256 ATGTGCTGGGACAGAGAGAGGATCTGGATTAGGCAACATCATAGAGTTGCTCTGAGCTGC 8315
QY 1256 TCTTTGGTGTAAACCTTCCAAATCTTAAATCTTTTGGAAATTCACAGCTCAAAGGAGGA 1315
DB 8316 TCTTTGGTGTAAACCTTCCAAATCTTAAATCTTTTGGAAATTCACAGCTCAAAGGAGGA 8375
QY 1316 AACCTACTCTCTGATCTACAGATGTTCTGCAATTTTCTATCATGCTCTATGGAACCTTC 1375

8376 AACCTACTCTCTGATCTACCAACATGTTCTGCAATTTTCTATCATGCTATGGAACCTC 8435
1376 TCTTAGAATCAGTGGCAAGATGTTCTATGATTAAGTGTCTGAGCTCAGGCAGGCA 1435
8436 TCTTAGAATCAGTGGCAAGATGTTCTATGATTAAGTGTCTGAGCTCAGGCAGGCA 8495
1436 GTCATGAATCTCTCTGAGTATTTTACTACTGATTTGCGGCGCAGCTCAGCTATCGGTT 1495
8496 GTCATGAATCTCTCTGAGTATTTTACTACTGATTTGCGGCGCAGCTCAGCTATCGGTT 8555
1496 TCTTACACCTCTCTATGAGAGTATCCATATTTATGTCGAGCGCAGTAAATGCTCCCCA 1555
8556 TCTTACACCTCTCTATGAGAGTATCCATATTTATGTCGAGCGCAGTAAATGCTCCCCA 8614
1556 CGAGATCAGTTTCTGAACCTAAGTGAATTTTATGAGGTTTATTTATGCACTATTA 1615
8615 CGAGATCAGTTTCTGAACCTAAGTGAATTTTATGAGGTTTATTTATGCACTATTA 8674
1616 AATCAACATTAAGTCTCTCTCTGATTTCTCTCTGATTAACATTAAGGCTCGCAAAAA 1675
8675 AATCAACATTAAGTCTCTCTCTGATTTCTCTCTGATTAACATTAAGGCTCGCAAAAA 8734
1676 AAAAAATCTTTTAAAAATAATGTCATAAATGTCATGTTCTCTGCGGCTTATGTCCTC 1735
8735 AAAAAATCTTTTAAAAATAATGTCATAAATGTCATGTTCTCTGCGGCTTATGTCCTC 8794
1736 TTTTCTTTTCTCTCTCTTTTCAACTAAGTCACGTCATTTTATTAAGATGGCCATAACTAT 1795
8795 TTTTCTTTTCTCTCTCTTTTCAACTAAGTCACGTCATTTTATTAAGATGGCCATAACTAT 8854
1796 TCAAAACCTATGCTGAGTTCCTCAAGCAGGCTCATAGTATGATCAAGTTGGATGGG 1855
8855 TCAAAACCTATGCTGAGTTCCTCAAGCAGGCTCATAGTATGATCAAGTTGGATGGG 8914
1856 CTACGGAGAACACGACCACTCTAGTTTATTTAAACCTGATTTTACTGCCACTTCC 1915
8915 CTACGGAGAACACGACCACTCTAGTTTATTTAAACCTGATTTTACTGCCACTTCC 8974
1916 CTTAGACTTGACCATATGACCCCTCGCT-CCCATTTCTAAGCATAGGGCGAGCTTTATT 1974
8975 CTTAGACTTGACCATATGACCCCTCGCT-CCCATTTCTAAGCATAGGGCGAGCTTTATT 9034
1975 TTTCAATGGTAAATA--GATATCATCTGAGTTTATCAAGAGTTGCGCGGGTGGTG 2031
9035 TTTCAATGGTAAATAAGATATCATCTGAGTTTATCAAGAGTTGCGCGGGTGGTG 9094
2032 AAAAGTTCAACACGATTCAGGTTTGTGTCGAGATTTCTAATTTTACATGTTTCTTT 2091
9095 AAAAGTTCAACACGATTCAGGTTTGTGTCGAGATTTCTAATTTTACATGTTTCTTT 9154
2092 TGCCAAAGGATTTTAAAAATAACATTTGTTTCTCTATCTGCTTTATGATGTC 2151
9155 TGCCAAAGGATTTTAAAAATAACATTTGTTTCTCTATCTGCTTTATGATGTC 9214
2152 GGAGACCATGAGAACACGCTCAATATCTTTTATGATCCCAAGCTGAAAGGCAAGCC 2211
9215 GGAGACCATGAGAACACGCTCAATATCTTTTATGATCCCAAGCTGAAAGGCAAGCC 9274
2212 CTCCAGAGAGGTTTATGTCACCAACACGAGCATTTGTCAGACCTTTGCGGGCTG 2271
9275 CTCCAGAGAGGTTTATGTCACCAACACGAGCATTTGTCAGACCTTTGCGGGCTG 9334
2272 TCTGAGCCATAGCTCTCCACGAGGACCT 2301.
9335 TCTGAGCCATAGCTCTCCACGAGGACCT 9364

RESULT 8

AAA35170

ID AAA35170 standard; DNA; 14879 BP.

XX

AC AAA35170;

XX

DT 28-JUL-2000 (first entry)
XX Human adenosine receptor related polynucleotide 2nd SEQ ID NO:44.
DE
XX
XX Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW phosphothioate; impaired respiration; inflammation; allergy;
KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW antiasthmatic; cytotstatic; analgesic; impaired airway;
KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX
XX Homo sapiens.
OS
XX WO200009525-A2.
PN
XX
XX 24-FEB-2000.
PD
XX 03-AUG-1999; 99WO-US017712.
PF
XX
XX 03-AUG-1998; 98US-0095212P.
PR
XX
XX (UYEC-) UNIV EAST CAROLINA.
PA
XX
XX Nyce JW;
PI
XX
XX WPI; 2000-205971/18.
DR
XX
XX New antisense oligonucleotides useful for treating e.g. pulmonary
PT vasoconstriction, inflammation, allergies, asthma, hypertension, or
PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT cancers.
XX
XX Disclosure; Page 1201-1205; 1343pp; English.
PS
XX
XX The present invention describes a new composition comprising an antisense
CC oligonucleotide (ON) with low adenosine (up to 15%), which targets
CC nucleic acids involved in bronchoconstriction, allergies, and/or
CC inflammation. The ON can have antiinflammatory, antiallergic,
CC antiasthmatic, cytostatic and analgesic activities. The compositions are
CC useful for the treatment of diseases associated with inflammation,
CC impaired airways, including lung disease and diseases whose secondary
CC effects afflict the lungs of a subject. They can be used for treating
CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma,
CC impaired respiration, respiratory distress syndrome, pain, cystic
CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
CC carcinomas, and cancers which may metastasise to the lungs, including
CC breast and prostate cancer. The reduction of the adenosine content of the
CC ONs reduces side effects. The A-containing ONs break down with the
CC release of deoxyadenosine which activates adenosine receptors causing
CC bronchoconstriction and inflammation. AAA3213 to AAA3512 represent the
CC nucleotide sequences given in the sequence listing from the present
CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185
CC sequences are also called SEQ ID NO:1 to 185, but the sequences differ
CC from the previously named sequences. SEQ ID NO:11 to 1680 (AAA3233 to
CC AAA3392) are specifically claimed ONs from the present invention. N.B.
CC Sequences given in the disclosure of the present invention do not match
CC up with their corresponding SEQ ID NO: sequences given in the sequence
CC listing
XX
SQ Sequence 14879 BP; 3971 A; 3370 C; 3434 G; 4103 T; 0 U; 1 Other;

Query Match

Best Local Similarity 97.4%; Score 2241.6; DB 3; Length 14879;

Matches 2294; Conservative 1; Mismatches 5; Indels 10; Gaps 4;

QY 1 CTGAATGATGCTCCCAAGTCTATGATGAGCTCTTGTGTCGCCAGTGAATAGGTG 60

DB 7056 CTGAATGATGCTCCCAAGTCTATGATGAGCTCTTGTGTCGCCAGTGAATAGGTG 7115

QY 61 TGTCATGTCATTTTAAAGACTATTAATACATAATATAGTTCTTCTCTCTTGG 120

Db 7116 TGTCCATGTGTCATTTTAAAGACTATTAAATACACTAATATAGTTCTCTCTCTTGG 7175
Qy 121 ATAATAGGACGTTGTTGCGTATGACCTTGGAGCCCTTAGTCCAAAGAGACCTTGGAGA 180
Db 7176 ATAATAGGACGTTGTTGCGTATGACCTTGGAGCCCTTAGTCCAAAGAGACCTTGGAGA 7235
Qy 181 ATTTTACTTCCCAAAAGGCAACAGACCCGTGAAATAGATGCAATGTCTAGGCCAAAG 240
Db 7236 ATTTTACTTCCCAAAAGGCAACAGACCCGTGAAATAGATGCAATGTCTAGGCCAAAG 7295
Qy 241 ACAAGAGTGTGGAAATTTTCCAAAGCAGGAAAGAACTCAGGTGACGAAACACGCTT 300
Db 7296 ACAAGAGTGTGGAAATTTTCCAAAGCAGGAAAGAACTCAGGTGACGAAACACGCTT 7355
Qy 301 TGTCTTTTCAATCAGTTTAAAGCCTCTGAACTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
Db 7356 TGTCTTTTCAATCAGTTTAAAGCCTCTGAACTCTCTCTCTCTCTCTCTCTCTCTCTCT 7415
Qy 361 TTTTAGAGAGACTAACAGAGACATTTGAAAGTCAGGTTAAAGCTGAATATAACATTGCTGA 420
Db 7416 TTTTAGAGAGACTAACAGAGACATTTGAAAGTCAGGTTAAAGCTGAATATAACATTGCTGA 7475
Qy 421 AATGTTTTCTTGTGTTATTTTAAAGGCTGAAGACATTTATGAGAAAGACTGGAATAA 480
Db 7476 AATGTTTTCTTGTGTTATTTTAAAGGCTGAAGACATTTATGAGAAAGACTGGAATAA 7535
Qy 481 TCATAAGAAAGAAAGACTGTTCCAAAGCTTGGGAAAGTGTATTTATCAGCAGTTAGT 540
Db 7536 TCATAAGAAAGAAAGACTGTTCCAAAGCTTGGGAAAGTGTATTTATCAGCAGTTAGT 7595
Qy 541 GAGAGAGAAAGAAATCAGAGAAAGTTTCAAGAGAACCTTAAGACAAACAGGTTAAGAGG 600
Db 7596 GAGAGAGAAAGAAATCAGAGAAAGTTTCAAGAGAACCTTAAGAGAACCAAGGTTAAGAGG 7655
Qy 601 AAGGAGAGAAATTAGTAAAGAGTTTCAAGAGAACCTTAGCCAGCTCAGTTAGTCCAG 660
Db 7656 AAGGAGAGAAATTAGTAAAGAGTTTCAAGAGAACCTTAGCCAGCTCAGTTAGTCCAG 7715
Qy 661 CAGCCTGTTCTCCAGCCCTTCTTACCAGGAGGTTAGAAAGCTTAGAAAAACAGTAGCAG 720
Db 7716 CAGCCTGTTCTCCAGCCCTTCTTACCAGGAGGTTAGAAAGCTTAGAAAAACAGTAGCAG 7775
Qy 721 AGGAGACTATGCTATGCTATAGATTAAGAGCAAGAGAAAGTCCCTCTTAATATTTCCA 780
Db 7776 AGGAGACTATGCTATGCTATAGATTAAGAGCAAGAGAAAGTCCCTCTTAATATTTCCA 7835
Qy 781 TGAAGCTCTGGAATGCAAAACCGATGCTCTGTACTTTTAGCACAATACATTTCTATC 840
Db 7836 TGAAGCTCTGGAATGCAAAACCGATGCTCTGTACTTTTAGCACAATACATTTCTATC 7895
Qy 841 AGGTAGATTTCCCAACCAAAATATATCAGAGATGCTTGTGATGCTGTTATATACAGC 900
Db 7896 AGGTAGATTTCCCAACCAAAATATATCAGAGATGCTTGTGATGCTGTTATATACAGC 7955
Qy 901 CTTTCCCTCTGAGTCAATGATTTTACACCTTCCCTGAGAAATCGAAAAATCAATTTGG 960
Db 7956 CTTTCCCTCTGAGTCAATGATTTTACACCTTCCCTGAGAAATCGAAAAATCAATTTGG 8015
Qy 961 GGAGCGGACATTTAGAAAGAAATCAAGTGTGATGATAATCAATTTCTTCAATAAGTT 1020
Db 8016 GGAGCGGACATTTAGAAAGAAATCAAGTGTGATGATAATCAATTTCTTCAATAAGTT 8075
Qy 1021 GCAGTTATTCAGATGCGCAAGGAAATAAAGTCAATTAGATAGGTTGGTAGAATTTAG 1080
Db 8076 GCAGTTATTCAGATGCGCAAGGAAATAAAGTCAATTAGATAGGTTGGTAGAATTTAG 8135
Qy 1081 AACATGCTGTTTTTTCAGGTTTATGCTC-----TTTTTTTTTTTTTTTTTTTTTAAATAGG 1135
Db 8136 AACATGCTGTTTTTTCAGGTTTATGCTC-----TTTTTTTTTTTTTTTTTTTTTAAATAGG 8195
Qy 1136 AATGCTGTTGTCAGAGCCCAATGCTATCCAAAAGCTCTCTCTTTCTCTGTCAGTC 1195

Db 8196 AATGCTGTTGTCAGAGCCCAATGCTATTTCCAAAAAGCTCTCTCTTTCTCTGTCAGTC 8255
Qy 1196 ATGTCCTGGGACAGAGAGGGATCTGGATTAGGCAACATCATAGAGTTGCTCTGAGCTGC 1255
Db 8256 ATGTCCTGGGACAGAGAGGGATCTGGATTAGGCAACATCATAGAGTTGCTCTGAGCTGC 8315
Qy 1256 TCTTTTGGTGATAAACCCCTTCCAAATCCTAAATCTTTTGGAAATTCACAAGCTCAAAAGGAGA 1315
Db 8316 TCTTTTGGTGATAAACCCCTTCCAAATCCTAAATCTTTTGGAAATTCACAAGCTCAAAAGGAGA 8375
Qy 1316 AACCTACTCTCTGATCTACACATGTTCTGCAATTTTCTATCATGTCATATGGAACCTTC 1375
Db 8376 AACCTACTCTCTGATCTACACATGTTCTGCAATTTTCTATCATGTCATATGGAACCTTC 8435
Qy 1376 TCTTAGAAATCCAGTGCAGAGAGTTCTATGATTAAAGTGTCTGAGCTCAGGCCAGGCA 1435
Db 8436 TCTTAGAAATCCAGTGCAGAGAGTTCTATGATTAAAGTGTCTGAGCTCAGGCCAGGCA 8495
Qy 1436 GTCATGAACCTCTCTGAGTTTACTACTGATTTGTGGGCGAGCCTCAGCTATCGGTT 1495
Db 8496 GTCATGAACCTCTCTGAGTTTACTACTGATTTGTGGGCGAGCCTCAGCTATCGGTT 8555
Qy 1496 TCTTCACACCTGTTATGAGAGTATCCATATTTATGTTTGGTGCAGG-CAGTAATGCTCCCCA 1555
Db 8556 TCTTCACACCTGTTATGAGAGTATCCATATTTATGTTTGGTGCAGG-CAGTAATGCTCCCCA 8614
Qy 1556 CGAGATCAGTTTCTGAACTAACCTGGAATTTTATGTTTGGTGTGTTTATATGCCCCAATATTA 1615
Db 8615 CGAGATCAGTTTCTGAACTAACCTGGAATTTTATGTTTGGTGTGTTTATATGCCCCAATATTA 8674
Qy 1616 AATCAACATTTACAGTTTCTTCCCTCTGATTTCTCCTGTAAACATTTAGCCTGCAAAAAA 1675
Db 8675 AATCAACATTTACAGTTTCTTCCCTCTGATTTCTCCTGTAAACATTTAGCCTGCAAAAAA 8734
Qy 1676 AAAAAATCTTTTAAAAATTAATGGCATAAAGTATTTGCTCTGGGCTACTGTATGCTTC 1735
Db 8735 AAAAAATCTTTTAAAAATTAATGGCATAAAGTATTTGCTCTGGGCTACTGTATGCTTC 8794
Qy 1736 TTTTCTTTTCTCTCTCTTTTCAACTAAGTCAACGCTCAATTTTAAAGTGGCCATAACTAT 1795
Db 8795 TTTTCTTTTCTCTCTCTTTTCAACTAAGTCAACGCTCAATTTTAAAGTGGCCATAACTAT 8854
Qy 1796 TCAAAACCTATGCTGAGTTCTCAAGCAGGCTCACAATAGTATGATGAAGTTGGGATGGGG 1855
Db 8855 TCAAAACCTATGCTGAGTTCTCTCAAGCAGGCTCGCATAGTATGATGAAGTTGGGATGGGG 8914
Qy 1856 CTACGGAAGAAACAGAACAACTCTAGTTTATTTAAACCTGTATTTACTGCCACTTCC 1915
Db 8915 CTACGGAAGAAACAGAACAACTCTAGTTTATTTAAACCTGTATTTACTGCCACTTCC 8974
Qy 1916 CCTTAGACTTGACCATATGACCCCTCGCT-CCCATTTCAAGCATAGGGCAGGCTTTAT 1974
Db 8975 CCTTAGACTTGACCATATGACCCCTTGTCCCATTTCAAGCATAGGGCAGGCTTTAT 9034
Qy 1975 TTTTACATGTTAATA--GATATCACTTCTGAGTTTATCAAAAGAGTTGGGGGGGGTGG 2031
Db 9035 TTTTACATGTTAATAAGATGATATCACTTGAAGTTTATCAAGAGTTGGGGGGGGTGG 9094
Qy 2032 AAAGTTCAACACAGATTCAGGTTTGTGTCAGATTTCTAAATTTTACATGTTCTTT 2091
Db 9095 AAAGTTCAACACAGATTCAGGTTTGTGTCAGATTTCTAAATTTTACATGTTCTTT 9154
Qy 2092 TGCCTAAAGGTTGATTTTTTAAATAAACATTTGTTTTCTCTTATCTTGTCTTTATTAGGTC 2151
Db 9155 TGCCTAAAGGTTGATTTTTTAAATAAACATTTGTTTTCTCTTATCTTGTCTTTATTAGGTC 9214
Qy 2152 GGAGACCATGAGAAAACAGGCTCAAAATCATCTTTTCTATGATCCCAAGCTGAAAGCAAGCC 2211
Db 9215 GGAGACCATGAGAAAACAGGCTCAAAATCATCTTTTCTATGATCCCAAGCTGAAAGCAAGCC 9274
Qy 2212 CTCAGAGAGGCTTATGTGACCCACAGACATTTGTCAGACACTTGGGGGCTG 2271
Db 9275 CTCAGAGAGGCTTATGTGACCCCAACCAAGCAATTTGTGACAGACTTTGGGGGCTG 9334

QY 2272 TCTGAAGCCATAGCTCCACGAGAGCCCT 2301
 |||||
 Db 9335 TCTGAAGCCATAGCTCCACGAGAGCCCT 9364

RESULT 9

AAF21292
 ID AAF21292 standard; DNA; 14879 BP.

XX AAF21292;

XX 14-MAR-2001 (first entry)

XX Human low adenosine antisense oligonucleotide related sequence #2859.

XX Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
 KW human; airway disorder; bronchoconstriction; lung inflammation;
 KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;
 KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
 KW respiratory obstruction; pulmonary obstruction; impeded respiration;
 KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
 KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
 KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
 KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
 KW cancer; ss.

XX Homo sapiens.

OS WO200062736-A2.

XX 26-OCT-2000.

XX 24-MAR-2000; 2000WO-US008020.

XX 06-APR-1999; 99US-0127958P.

XX (UYEC-) UNIV EAST CAROLINA.

XX (NYCE/) NYCE J W.

PI Nyce JW;

XX WPI; 2000-679539/66.

XX Low adenosine (A) content antisense oligonucleotides which do not trigger
 PT adenosine receptors during metabolism, useful e.g. for treating cancers
 PT and respiratory obstructions.

PS Disclosure; Page 1284-1288; 1592pp; English.

XX The present invention describes low adenosine (A) content antisense
 CC oligonucleotides and compositions (I) comprising them. In the antisense
 CC oligonucleotides the A is replaced by a 'Universal' or alternative base.
 CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
 CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
 CC The antisense oligonucleotides and (I) can be used to down-regulate the
 CC expression and or activity of target polypeptides associated with
 CC lung/respiratory disorders and malignancies, such as stimulating and
 CC activating peptide factors and transmitters, transcription factors,
 CC immunoglobulins and antibodies, antibody receptors, cytokines and
 CC chemokines, endogenously produced specific and non-specific enzymes,
 CC binding proteins, adhesion molecules and their receptors, cytokine and
 CC chemokine receptors, adenosine receptors, bradykinin receptors, central
 CC nervous system (CNS) and peripheral nervous and non-nervous system
 CC receptors, CNS and peripheral nervous and non-nervous system peptide
 CC transmitters, defensins, growth factors, vasoactive peptides and
 CC antisense oligonucleotides may be used in this way to treat disorders
 CC including respiratory obstruction (especially pulmonary obstruction
 CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or
 CC surfactant hypoproduction which are associated with a disease or
 CC condition selected from pulmonary vasoconstriction, inflammation,
 CC allergies, asthma, impeded respiration, respiratory distress syndrome

CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
 CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
 CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
 CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
 CC fragments and antisense oligonucleotides used in the exemplification of
 CC the present invention

XX Sequence 14879 BP; 3971 A; 3370 C; 3434 G; 4103 T; 0 U; 1 Other;

Query Match 97.4%; Score 2241.6; DB 3; Length 14879;
 Best Local Similarity 99.3%; Pred. NO. 0;
 Matches 2294; Conservative 1; Mismatches 5; Indels 10; Gaps 4;

QY	1	CTGAATGATGCTCCCAAGTCTATGTGATGAGCTCTTGTGTCGCCAGTGAATAGTG	60
Db	7056	CTGAATGATGCTCCCAAGTCTATGTGATGAGCTCTTGTGTCGCCAGTGAATAGTG	7115
QY	61	TGTCCATGTGTCATTTTAAAGACTATTAACTAATAATATAGTTTCTTCTCTTGG	120
Db	7116	TGTCCATGTGTCATTTTAAAGACTATTAACTAATAATATAGTTTCTTCTCTTGG	7175
QY	121	ATAATAGGACGCTTCTCGTATGACCTGGAGCCCTAGCTCCAAAGAGCCCTGGAGA	180
Db	7176	ATAATAGGACGCTTCTCGTATGACCTGGAGCCCTAGCTCCAAAGAGCCCTGGAGA	7235
QY	181	ATTTACTTCCCAAGGCAACAGACCCGTGAAATAGATGCAATGTCTAGCCAAAAG	240
Db	7236	ATTTACTTCCCAAGGCAACAGACCCGTGAAATAGATGCAATGTCTAGCCAAAAG	7295
QY	241	ACAAGAAGTGCTGGAAATTTTGCAGGCAAGAAAGAACTCAGGTGACGAGAACACCTT	300
Db	7296	ACAAGAAGTGCTGGAAATTTTGCAGGCAAGAAAGAACTCAGGTGACGAGAACACCTT	7355
QY	301	TGCTTTTCAATCAGTTTAAAGCCCTCCTGAACTCCTCTCTCATGTGCTACTGCCCT	360
Db	7356	TGCTTTTCAATCAGTTTAAAGCCCTCCTGAACTCCTCTCTCATGTGCTACTGCCCT	7415
QY	361	TTTTAGAGAGACTAAACAGAGACATTTGAAAGTCAGGTAAGCTGAATATACATTCCTGA	420
Db	7416	TTTTAGAGAGACTAAACAGAGACATTTGAAAGTCAGGTAAGCTGAATATACATTCCTGA	7475
QY	421	AATGTTTTCCTGTGTATTTTAAAGGCTGAAACACATTTAGGAAAGACTGGAATAA	480
Db	7476	AATGTTTTCCTGTGTATTTTAAAGGCTGAAACACATTTAGGAAAGACTGGAATAA	7535
QY	481	TCATAGAAAGGAAAGAGACTGTTCCAAAGCTTGGGAAAAAGTGTATTTATCAGCAGTTAGT	540
Db	7536	TCATAGAAAGGAAAGAGACTGTTCCAAAGCTTGGGAAAAAGTGTATTTATCAGCAGTTAGT	7595
QY	541	GAGAGGAGAAAAATCAGAGAGAGTTTCAAGAGGCTGAAACACATTTAGGAAAGACTGGAATAA	600
Db	7596	GAGAGGAGAAAAATCAGAGAGAGTTTCAAGAGGCTGAAACACATTTAGGAAAGACTGGAATAA	7655
QY	601	AAGGAGAAAAATTAGGTAAGAGGTTTCAAGAGCACTAGCCCGCAGTCAGTGATGCCAG	660
Db	7656	AAGGAGAAAAATTAGGTAAGAGGTTTCAAGAGCACTAGCCCGCAGTCAGTGATGCCAG	7715
QY	661	CAGCCTGTCTCCAGCCCTTCTTACCGGCGAGGTGAAAGACTTTAGAAACAGTAGCAG	720
Db	7716	CAGCCTGTCTCCAGCCCTTCTTACCGGCGAGGTGAAAGACTTTAGAAACAGTAGCAG	7775
QY	721	AGGAGATCTATCATCTTATAGATTAAAGAGGCAAAAAAGAAATCCCTCTTAAATATTTCCA	780
Db	7776	AGGAGATCTATCATCTTATAGATTAAAGAGGCAAAAAAGAAATCCCTCTTAAATATTTCCA	7835
QY	781	TGAAGCTCTGGAATGCAACCGATGCTCTCTGTACTTTTACGACATACCATTTCACTAC	840
Db	7836	TGAAGCTCTGGAATGCAACCGATGCTCTCTGTACTTTTACGACATACCATTTCACTAC	7895
QY	841	AGGTAGATTTTCCCAACCAAAATATATCCAGAGATGCCCTTTGTTCATTGGGTTATATACAGC	900
Db	7896	AGGTAGATTTTCCCAACCAAAATATATCCAGAGATGCCCTTTGTTCATTGGGTTATATACAGC	7955

QY 901 CTTTGGCTCTCTGAGTCAATGATTTACCACTTTCCCTGAGAAATCGAAATCAATTTGG 960
DB 7956 CTTTGGCTCTCTGAGTCAATGATTTACCACTTTCCCTGAGAAATCGAAATCAATTTGG 8015
QY 961 GGAGCGGACATTTAGAAAAAGAAATCAAAAGTGTCTCATGATATCAAAATCTTCAATAAGTT 1020
DB 8016 GGAGCGGACATTTAGAAAAAGAAATCAAAAGTGTCTCATGATATCAAAATCTTCAATAAGTT 8075
QY 1021 CGAGTTATTCAGATGGCCAAAGGAAAAATAAAGTCAATAGATAGGTTGGTAGAATTTAG 1080
DB 8076 CGAGTTATTCAGATGGCCAAAGGAAAAATAAAGTCAATAGATAGGTTGGTAGAATTTAG 8135
QY 1081 AACATGCTGTTTTCAGGTTTATGATC-----TTTTTTTTTTTTTTTTTTTTTTTAAATAGGG 1135
DB 8136 AACATGCTGTTTTCAGGTTTATGATC-----TTTTTTTTTTTTTTTTTTTTTTTAAATAGGG 8195
QY 1136 AAATGTGTTTGGTGAGAGCAATGTCAATCAAAAAGCTCTCTTTTCTGTGTCTAGTC 1195
DB 8196 AAATGTGTTTGGTGAGAGCAATGTCAATCAAAAAGCTCTCTTTTCTGTGTCTAGTC 8255
QY 1196 ATGCTCGGAGAGAGGAGATCGGATTAAGGCAACATCATAGAGTTGCTCTCAGCTGC 1255
DB 8256 ATGCTCGGAGAGAGGAGATCGGATTAAGGCAACATCATAGAGTTGCTCTCAGCTGC 8315
QY 1256 TCTTTGGTGATAACCTTTCCAAATCTAAACTTTTGGAAATTCACAGCTCAAAAGGAGGA 1315
DB 8316 TCTTTGGTGATAACCTTTCCAAATCTAAACTTTTGGAAATTCACAGCTCAAAAGGAGGA 8375
QY 1316 AACCTACTCTGATCTACACATGTTCTGCAATTTTCTATCATGTTCTATGGAACCTC 1375
DB 8376 AACCTACTCTGATCTACACATGTTCTGCAATTTTCTATCATGTTCTATGGAACCTC 8435
QY 1376 TCTTAGAATCCAGTGGCAAGAGTTCTATGATTAAGTGTCTGAGCTCAGGCCAGGCA 1435
DB 8436 TCTTAGAATCCAGTGGCAAGAGTTCTATGATTAAGTGTCTGAGCTCAGGCCAGGCA 8495
QY 1436 GTATGAACTACTCTGAGTTATTTACTACTGATTTGTGGGGAGCTCAGCTATCGTT 1495
DB 8496 GTATGAACTACTCTGAGTTATTTACTACTGATTTGTGGGGAGCTCAGCTATCGTT 8555
QY 1496 TCTTCACACCTGCTTATGAGATATCCATATTTATGTTGTCGAGGCGCAATGCTCCCA 1555
DB 8556 TCTTCACACCTGCTTATGAGATATCCATATTTATGTTGTCGAGGCGCAATGCTCCCA 8614
QY 1556 CGAGATCAGTTCTGAACTAACTGGAATTTTATGAGTTTATGAGTCAACTATTA 1615
DB 8615 CGAGATCAGTTCTGAACTAACTGGAATTTTATGAGTTTATGAGTCAACTATTA 8674
QY 1616 AATCAACATTTACAGTTCTTCCCTCTGATTTCTCCTGTAAACATTAAGGCTGCAAAAA 1675
DB 8675 AATCAACATTTACAGTTCTTCCCTCTGATTTCTCCTGTAAACATTAAGGCTGCAAAAA 8734
QY 1676 AAAAAATCTTTTAAAAATAATGCAATAAAGTATTTGCTCTGGGCTTACTGTATGCTC 1735
DB 8735 AAAAAATCTTTTAAAAATAATGCAATAAAGTATTTGCTCTGGGCTTACTGTATGCTC 8794
QY 1736 TTTTCTTTTCTCTCTTTTCAACTAAGTCAAGTCAATTTATTAAGTGGCCATAACTAT 1795
DB 8795 TTTTCTTTTCTCTCTTTTCAACTAAGTCAAGTCAATTTATTAAGTGGCCATAACTAT 8854
QY 1796 TCAAAACCTATGCTGAGTTCTTCAAGGAGGAGTCAATAGTGAAGTTGGGATGGG 1855
DB 8855 TCAAAACCTATGCTGAGTTCTTCAAGGAGGAGTCAATAGTGAAGTTGGGATGGG 8914
QY 1856 CTACGGAGAAACAGAACTCTAGTTTATTTAAAACTGTATTTACTGCCCACTTCC 1915
DB 8915 CTACGGAGAAACAGAACTCTAGTTTATTTAAAACTGTATTTACTGCCCACTTCC 8974
QY 1916 CTTTAGACTTGAACATATGACCCCTCGCT-CCCATTTCAAGCATAGGGGAGGCTTATT 1974
DB 8975 CTTTAGACTTGAACATATGACCCCTCGCT-CCCATTTCAAGCATAGGGGAGGCTTATT 9034
QY 1975 TTTCAATAGTGAATA---GATATCACTTGAGGTTTATCAAAAGATTCGGCGGGTGGTG 2031

DB 9035 TTTCACATGGTATAGATGATATCATCTGAGGTTTATCAAGAGTTGCGGCGGGTGGTG 9094
QY 2032 AAAGTTCAACAACAGATTCAGGTTTGTGTGGCCAGATTTAAATTTACATGTTCTTT 2091
DB 9095 AAAGTTCAACAACAGATTCAGGTTTGTGTGGCCAGATTTAAATTTACATGTTCTTT 9154
QY 2092 TGCCAAAGGTTGATTTTAAATAAATTTGTTTCTCTTATCTTCTTATTTAGGTC 2151
DB 9155 TGCCAAAGGTTGATTTTAAATAAATTTGTTTCTTATCTTATCTTGTATTTAGGTC 9214
QY 2152 GGAGACCATGAGAAACAGGCTCAAAATCATCTTTTATGATCCCAAGCTCAAAAGGCAAGCC 2211
DB 9215 GGAGACCATGAGAAACAGGCTCAAAATCATCTTTTATGATCCCAAGCTCAAAAGGCAAGCC 9274
QY 2212 CTTCAGAGAGCTTATGTACCCACACAGGAGCAATTTGGTGCACAGCTTCGGGGCTG 2271
DB 9275 CTTCAGAGAGCTTATGTACCCACACAGGAGCAATTTGGTGCACAGCTTCGGGGCTG 9334
QY 2272 TCTGAAGCCATAGCTTCCACGGAGAGCCCT 2301
DB 9335 TCTGAAGCCATAGCTTCCACGGAGAGCCCT 9364

RESULT 10
ABZ96986
ID ABZ96986 standard; DNA; 14879 BP.
XX
AC ABZ96986;
XX
DT 17-OCT-2003 (first entry)
XX
DE Human nucleic acid sequence.
XX
KW Human; antisense; lung dysfunction; nasal airway dysfunction;
KW antiinflammatory steroid; ubiquinone; antiinflammatory; antiallergic;
KW antisthmatic; hypotensive; immunosuppressive; cytostatic; gene therapy;
KW antisense gene therapy; respiratory; lung; adenosine sensitivity;
KW adenosine receptor; bronchodilation; bronchoconstriction; lung allergy;
KW lung inflammation; respiratory disease; ds.
XX
OS Homo sapiens.
XX
PN WO200285308-A2.
XX
PD 31-OCT-2002.
XX
PF 23-APR-2002; 2002WO-US013135.
XX
PR 24-APR-2001; 2001US-0286137P.
XX
PA (BPIG-) EPIGENESIS PHARM INC.
XX
PI Nyce JW, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D;
PI Miller S, Tang L, Shahabuddin S;
XX
WP1; 2003-229219/22.
XX
PT Pharmaceutical composition for treating ailments associated with impaired
PT respiration, has oligo(s) antisense to specific gene(s) or its
PT corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or
PT ubiquinone.
XX
PS Disclosure; SEQ ID NO 12228; 872pp; English.
XX
CC The invention relates to a novel pharmaceutical composition, which has a
CC first active agent comprising an oligonucleotide antisense to the
CC initiation codon, coding region, 5' or 3' end genomic flanking regions,
CC 5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of
CC junctions of genes encoding a polypeptide associated with lung and/or
CC nasal airway dysfunction and a second active agent comprising an
CC antiinflammatory steroid and ubiquinone. A composition of the invention
CC has antiinflammatory, antiallergic, antiasthmatic, hypotensive,

CC immunosuppressive, and cytostatic activity. The composition may have a
CC use in antisense gene therapy. The composition is useful for treating or
CC preventing a respiratory, lung or malignant disease or condition, also
CC for enhancing the prophylactic or therapeutic respiratory effect of an
CC antiinflammatory steroid in a subject, for reducing or depleting levels
CC of, or reducing sensitivity to adenosine, reducing levels of adenosine
CC receptor, producing bronchodilation, increasing levels of ubiquinone or
CC lung surfactant in a subject's tissue, or treating bronchoconstriction,
CC lung inflammation, lung allergies, or a respiratory disease or condition.
CC Note: the sequence data for this patent is not represented in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 14879 BP; 3971 A; 3370 C; 3434 G; 4103 T; 0 U; 1 Other;

Query Match 97.4%; Score 2241.6; DB 7; Length 14879;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2294; Conservative 1; Mismatches 5; Indels 10; Gaps 4;

QY 1 CTGAATGATGCTCCCAAGTGTATGTGATGAGCTCCTTGTTGTCGCCAGTGGATAGTG 60
DB |||||
QY 61 TGTCATGTGCTATTTAAAGACTATTAATPACACTAATATAGTTTCTTCTCTTTGG 120
DB |||||
QY 121 ATATAGCAGCTGTTCCGATGAGCTTGAAGCCCTAGTCCCAAGAGCCTTTGAGA 180
DB |||||
QY 1716 ATATAGCAGCTGTTCCGATGAGCTTGAAGCCCTAGTCCCAAGAGCCTTTGAGA 7235
DB |||||
QY 181 ATTTACTTCCCAAGGCNA CAGACCGTGAATAATAGATGCCAATGTGTCAGCAAAAAG 240
DB |||||
QY 7236 ATTTACTTCCCAAGGCNA CAGACCGTGAATAATAGATGCCAATGTGTCAGCAAAAAG 7295
DB |||||
QY 241 ACAGAGTGTGGAATTTTGCACGAGCAAGAAAGACTCAGGTGAGCAACACCTT 300
DB |||||
QY 7296 ACAGAGTGTGGAATTTTGCACGAGCAAGAAAGACTCAGGTGAGCAACACCTT 7355
DB |||||
QY 301 TGCCTTTCAATCAGTTTAAACAGCTCCTGAACTCCTTCTATCATGTGTCCTTCCCTG 360
DB |||||
QY 7356 TGCCTTTCAATCAGTTTAAACAGCTCCTGAACTCCTTCTATCATGTGTCCTTCCCTG 7415
DB |||||
QY 361 TTTTAGAGAGACTAACAGAGACATTTAAAGTCAAGGTAAGCTCAATATACATTTCTGA 420
DB |||||
QY 7416 TTTTAGAGAGACTAACAGAGACATTTAAAGTCAAGGTAAGCTCAATATACATTTCTGA 7475
DB |||||
QY 421 AATGTTTTCTCTGTGATTTTAAACAGGCTGAAGACATTTAGAGAAAGACTGGAAATA 480
DB |||||
QY 7476 AATGTTTTCTCTGTGATTTTAAACAGGCTGAAGACATTTAGAGAAAGACTGGAATA 7535
DB |||||
QY 481 TCATAAGAAAGAAAGACTCTTCAAGCTTGGGAAAGATGATTTATTCAGAGTTAGT 540
DB |||||
QY 7536 TCATAAGAAAGAAAGACTCTTCAAGCTTGGGAAAGATGATTTATTCAGAGTTAGT 7595
DB |||||
QY 541 GAGAGGAAGAAATTCAGAGAGATTCAGAGAGACACTAAGACAAACAGGTAAAGAGG 600
DB |||||
QY 7596 GAGAGGAAGAAATTCAGAGAGATTCAGAGAGACACTAAGACAAACAGGTAAAGAGG 7655
DB |||||
QY 601 AAGGAAGAAATTCAGAGAGATTCAGAGAGACACTAAGACAAACAGGTAAAGAGG 660
DB |||||
QY 7656 AAGGAAGAAATTCAGAGAGATTCAGAGAGACACTAAGACAAACAGGTAAAGAGG 7715
DB |||||
QY 661 CAGCTGTTCTCCAGCCCTTCTTACCCGCGGCTGAAGACTTACAGAAACAGTACAG 720
DB |||||
QY 7716 CAGCTGTTCTCCAGCCCTTCTTACCCGCGGCTGAAGACTTACAGAAACAGTACAG 7775
DB |||||
QY 721 AGGAGATCTATGCTCTATAGATTTAAAGAGCAAAAGAAATCCCTCTTAAATATTTCCA 780
DB |||||
QY 7776 AGGAGATCTATGCTCTATAGATTTAAAGAGCAAAAGAAATCCCTCTTAAATATTTCCA 7835
DB |||||
QY 781 TGAAGCTCTGAAGTGAACCGATGCTCTGTTACTTTTAGCACATACCACTTTCATCTAC 840
DB |||||

DB 7836 TGAAGCTCTGGAATGCAAAACCGATCTCCTCTGTACCTTTAGCACATACCAATTCATCTAC 7895
QY 841 AGGTAGATTTCCCAACCAAAATATATCCAGAGATGCCCTTTGTCTATGCGGTTATATACAGC 900
DB 7896 AGGTAGATTTCCCAACCAAAATATATCCAGAGATGCCCTTTGTCTATGCGGTTATATACAGC 7955
QY 901 CTTTGCCTCTGTAGTCAATGATTTTACACATTTCCCTGAGAAATCGAAATCATTTTGG 960
DB 7956 CTTTGCCTCTGTAGTCAATGATTTTACACATTTCCCTGAGAAATCGAAATCATTTTGG 8015
QY 961 GGAGCGGACATTTAGAAAAGAAATCAAAAGTGTGATGGAATCAAAATCTTCAATAGTT 1020
DB 8016 GGAGCGGACATTTAGAAAAGAAATCAAAAGTGTGATGGAATCAAAATCTTCAATAGTT 8075
QY 1021 GCAGTTATTCAGATGCCCAAGGAAATTAAGTCAATAGATAGGTTGGTGTAGATTTAG 1080
DB 8076 GCAGTTATTCAGATGCCCAAGGAAATTAAGTCAATAGATAGGTTGGTGTAGATTTAG 8135
QY 1081 AACATGCTCTTTTTCAGGTTTATGTC-----TTTTTTTTTTTTTTTTTTTTTAAATAGG 1135
DB 8136 AACATGCTCTTTTTCAGGTTTATGTC-----TTTTTTTTTTTTTTTTTTTTTAAATAGG 8195
QY 1136 AAATGTTTTGTGAGAGCCCAATGTCAATCCAAAAGCTCTCTTTTCTGCTGCTCAGTC 1195
DB 8196 AAATGTTTTGTGAGAGCCCAATGTCAATCCAAAAGCTCTCTTTTCTGCTGCTCAGTC 8255
QY 1196 ATGTGCTGGACAGAGAGGATCTGGATTAGGCAACATCATAGATTTGCTCTGAGCTGC 1255
DB 8256 ATGTGCTGGACAGAGAGGATCTGGATTAGGCAACATCATAGATTTGCTCTGAGCTGC 8315
QY 1256 TCTTTGGTGATTAACCTTCCAAATCCTAAAATTTTGGAAATTCACAGCTCAAGAGGAG 1315
DB 8316 TCTTTGGTGATTAACCTTCCAAATCCTAAAATTTTGGAAATTCACAGCTCAAGAGGAG 8375
QY 1316 AACCTACTCTCTGATCTACCATCTTCTGCAATTTTCTATCATGCTCTATGGAACCTTC 1375
DB 8376 AACCTACTCTCTGATCTACCATCTTCTGCAATTTTCTATCATGCTCTATGGAACCTTC 8435
QY 1376 TCTTAGAAATCCAGTGGCAAGATCTTATGATTAAGTGTCTGAGCTCAAGGCAAGCA 1435
DB 8436 TCTTAGAAATCCAGTGGCAAGATCTTATGATTAAGTGTCTGAGCTCAAGGCAAGCA 8495
QY 1436 GTCTAGAACTACTCTGAGTTATTTACTACTGATTTGTTGGGCGAGCTCAGCTATCGTT 1495
DB 8496 GTCTAGAACTACTCTGAGTTATTTACTACTGATTTGTTGGGCGAGCTCAGCTATCGTT 8555
QY 1496 TCTTCAACCTCTGCTTATGAGAGTATCCATATTTATGTTGCGAGGCGAGTAAATGCTCCCA 1555
DB 8556 TCTTCAACCTCTGCTTATGAGAGTATCCATATTTATGTTGCGAGG-CAGTAATGCTCCCA 8614
QY 1556 CGAGATCAGTTTCTCAACTAACCTGGAATTTTATGGGTTTTTATGTCGCACTATTA 1615
DB 8615 CGAGATCAGTTTCTCAACTAACCTGGAATTTTATGGGTTTTTATGTCGCACTATTA 8674
QY 1616 AATCAACATTTACAGTTCTTCCCTCTGATTTCTCTCTGTAATTTCTCTGTAACATTA 1675
DB 8675 AATCAACATTTACAGTTCTTCCCTCTGATTTCTCTCTGTAACATTAAGGCTGCAAAAA 8734
QY 1676 AAAAAATCTTTTAAAAATTAATGTCATTAAGTATTTGCTCTGGGCTTACTGTATGCTTC 1735
DB 8735 AAAAAATCTTTTAAAAATTAATGTCATTAAGTATTTGCTCTGGGCTTACTGTATGCTTC 8794
QY 1736 TTTTCTTTTCTCTCTTTTCAACTAAGTCAAGTCAATTTATTAAGATGGCCATAACTAT 1795
DB 8795 TTTTCTTTTCTCTCTTTTCAACTAAGTCAAGTCAATTTATTAAGATGGCCATAACTAT 8854
QY 1796 TCAAAACCTATGCTGAGTTTCTCAAGGAGGCTCAGATGATGATGAGGTTGGGATGGG 1855
DB 8855 TCAAAACCTATGCTGAGTTTCTCAAGGAGGCTCAGATGATGATGAGGTTGGGATGGG 8914
QY 1856 CTACGGAAGAAACCAAGCAAACTCTAGTTTATTTAAACCTGTTTACTGCCCACTTCC 1915
DB 8915 CTACGGAAGAAACCAAGCAAACTCTAGTTTATTTAAACCTGTTTACTGCCCACTTCC 8974

QY 1916 CCTTAGACTTGACCATATGACCCCTCGCT-CCATTCTAAGCATAGGGCAGGCTTTATT 1974
DB 8975 CCTTAGACTTGACCATATGACCCCTTGCTCCCATCTAAGCATAGGGCAGGCTTTATT 9034
QY 1975 TTTACAATGGTAATA--GATATCACTTGAGGTTTATCAAGAGTTGCGGCGGTGGTG 2031
DB 9035 TTTACAATGGTAATAGATGATCACTTGAGGTTTATCAAGAGTTGCGGCGGTGGTG 9094
QY 2032 AAAGTTCAACAACGAGTTCAGTTTGGTTTGGCCAGATCTTAATTTACATGTTTCTTT 2091
DB 9095 AAAGTTCAACAACGAGTTCAGTTTGGTTTGGCCAGATCTTAATTTACATGTTTCTTT 9154
QY 2092 TGCCAAAGGTCATTTTAAATAAACAATTTGTTTCTTATCTTCTTATTTAGTGC 2151
DB 9155 TGCCAAAGGTCATTTTAAATAAACAATTTGTTTCTTATCTTCTTATTTAGTGC 9214
QY 2152 GGAGACCATGAGAAACAGCGTCAATCATCTTTTCATGATCCCAAGCTGAAAGCAAGCC 2211
DB 9215 GGAGACCATGAGAAACAGCGTCAATCATCTTTTCATGATCCCAAGCTGAAAGCAAGCC 9274
QY 2212 CTCGAGAGCGTTATGTGACCCACACCGAGCACATTTGGTGACAGACTTCGGGGCCTG 2271
DB 9275 CTCGAGAGCGTTATGTGACCCACACCGAGCACATTTGGTGACAGACTTCGGGGCCTG 9334
QY 2272 TCTGAAGCCATAGCCTCCACGGAGAGCCCT 2301
DB 9335 TCTGAAGCCATAGCCTCCACGGAGAGCCCT 9364
RESULT 11
AAD28049
ID AAD28049 standard; DNA; 12459 Bp.
XX
AC AAD28049;
XX
DT 09-AUG-2002 (first entry)
XX
DE Human endothelin-1 gene.
XX
KW Single nucleotide polymorphism; SNP; endothelin-1; EDN-1; hypertension;
end stage renal disease; ESRD; non-insulin dependent diabetes mellitus;
NIDDM; cancer; lung; breast; prostate; colon; cerebrovascular accident;
atherosclerotic peripheral vascular disease; cataract; cardiomyopathy;
myocardial infarction; atrial fibrillation; alcohol abuse; anxiety;
alcoholic cirrhosis; asthma; chronic obstructive pulmonary disease; COPD;
cholecystectomy; degenerative joint disease; seizure disorder; human; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT repeat_region 98..383
FT protein_bind /*tag= a
FT /*tag= b
FT /*tag= c
FT misc_feature 979..1039
FT /*tag= c
FT protein_bind /*note= "putative Z-DNA region"
FT /*tag= d
FT variation /*bound_moiety= "Acute phase reactant regulatory element"
FT /*tag= e
FT /*standard_name= "Single nucleotide polymorphism"
FT variation /*tag= f
FT /*standard_name= "Single nucleotide polymorphism"
FT protein_bind 2951..2958
FT /*tag= g
FT protein_bind /*bound_moiety= "TPA/JUN"
FT /*tag= h

protein_bind /bound_moiety= "TPA/JUN"
FT 3316..3328
FT /*tag= i
protein_bind /bound_moiety= "TPA/JUN"
FT 3499..3505
FT /*tag= j
CAAT_signal /bound_moiety= "TPA/JUN"
FT 3510..3515
FT /*tag= k
TATA_signal 3577..3582
FT /*tag= l
CDS 3608..10441
FT /*tag= n
/product= "Protein encoded by endothelin-1 gene"
/transl_except= (pos:3755..3763, aa:Gly-Arg)
/transl_except= (pos:3794..3802, aa:Leu-Gly)
/transl_except= (pos:3803..3811, aa:Thr-Ser)
/transl_except= (pos:5588..5596, aa:Ser-Ala)
/transl_except= (pos:5690..5698, aa:Pro-Trp)
/transl_except= (pos:7210..7218, aa:Pro-Val)
/transl_except= (pos:7261..7275, aa:Pro-Glu-Met)
/transl_except= (pos:7282..7290, aa:Cys-Pro)
/transl_except= (pos:7503..7511, aa:Gly-Arg)
/transl_except= (pos:7530..7544, aa:Glu-Ser-Glu)
/transl_except= (pos:9248..9256, aa:Ser-Ser)
/transl_except= (pos:9237..9245, aa:Phe-Gly)
/transl_except= (pos:9821..9829, aa:Leu-Asn)
/transl_except= (pos:9929..9937, aa:Ile-Leu)
/transl_except= (pos:9995..10003, aa:Ser-Pro)
/transl_except= (pos:10016..10024, aa:Val-Cys)
/transl_except= (pos:10058..10072, aa:Tyr-Leu-Ser)
/transl_except= (pos:10112..10120, aa:Pro-Phe)
/transl_except= (pos:10157..10165, aa:Tyr-Lys)
/transl_except= (pos:10190..10198, aa:Gln-Asn)
/transl_except= (pos:10262..10270, aa:Leu-Ile)
/transl_except= (pos:10345..10354, aa:Phe-Asn)
/transl_except= (pos:10361..10369, aa:Phe-met)
/note= "Does not include start and stop codons. The CDS
includes 5 exons and 4 introns. The translation start
site is noted as being at position 3876 of the sequence
(SQ ID NO: 1) throughout the specification. However, the
translation is shown to start at position 3608 of the
sequence in the sequence listing"
exon /partial
FT 3608..3939
FT /*tag= o
misc_signal 3608
FT /*tag= m
/note= "transcription start site"
FT 3876..3938
FT /*tag= p
FT intron 3940..5594
FT /*tag= q
FT protein_bind 4612..4618
FT /*tag= r
FT /*bound_moiety= "Unspecified"
FT 5585..5753
FT /*tag= s
FT intron 5754..7182
FT /*tag= t
FT exon 7183..7338
FT /*tag= u
FT intron 7339..7502
FT /*tag= v
FT exon 7503..7646
FT /*tag= w
FT intron 7647..9211
FT /*tag= x
FT protein_bind 8205..8218
FT /*tag= y
FT /*bound_moiety= "Unspecified"
FT 8260..8266
FT /*tag= z

```
FT exon /bound_moiety= "Unspecified"
FT 9212..10443
FT *tag= aa
FT protein_bind 9392..9398
FT *tag= ab
FT /bound_moiety= "Unspecified"
FT protein_bind 10751..10764
FT *tag= ac
FT /bound_moiety= "Unspecified"
XX
XX WO200222881-A1.
XX
XX 21-MAR-2002.
XX
XX 11-SEP-2001; 2001WO-US028834.
XX
XX 11-SEP-2000; 2000US-0231672P.
XX 27-OCT-2000; 2000US-0243814P.
XX
XX (DZGE-) DZGENES LLC.
XX
XX Moskowitz DW;
XX
XX WPI; 2002-393976/42.
XX P-PSDB; AAE14677.
XX
XX Diagnosing genetic susceptibility for disease in a subject, by obtaining
XX a sample containing nucleic acid and detecting the presence/absence of a
XX single nucleotide polymorphism in the endothelin-1 gene.
XX
XX Claim 7; Page 72-81; 84pp; English.
XX
XX The invention relates to a method of diagnosing a genetic susceptibility
XX for a disease in a subject, involving obtaining a biological sample
XX containing nucleic acid from the subject, and analysing the nucleic acid
XX to detect a single nucleotide polymorphism (SNP) in the endothelin-1 (EDN
XX -1) gene, where the SNP is associated with a genetic predisposition for a
XX disease or a disorder. Examples of disorders include hypertension (HTN),
XX end stage renal disease (ESRD) due to HTN, non-insulin dependent diabetes
XX mellitus (NIDDM), ESRD due to NIDDM, lung cancer, breast cancer, prostate
XX cancer, colon cancer, atherosclerotic peripheral vascular disease (ASPWD)
XX due to HTN or NIDDM, cerebrovascular accident (CVA) due to HTN or NIDDM,
XX cataracts due to HTN, cardiomyopathy (CM) with HTN, myocardial infarction
XX (MI) due to HTN or NIDDM, ischaemic cardiomyopathy, ischaemic
XX cardiomyopathy with NIDDM, atrial fibrillation (afib) without valvular
XX disease, alcohol abuse, alcoholic cirrhosis, anxiety, asthma, chronic
XX obstructive pulmonary disease (COPD), cholecystectomy, degenerative joint
XX disease (DJD), ESRD and frequent de-clots, ESRD due to focal segmental
XX glomerular sclerosis (FSGS), and ESRD due to insulin dependent diabetes
XX mellitus (IDDM) or seizure disorder. The invention also provides a method
XX for treatment or prophylaxis of diseases associated with SNPs in
XX endothelin-1 gene. The present sequence is human endothelin-1 gene
XX containing SNPs in the promoter region
XX
XX Sequence 12459 BP; 3372 A; 2727 C; 2816 G; 3543 T; 0 U; 1 Other;
XX
XX Query Match 97.3%; Score 2239.6; DB 6; Length 12459;
XX Best Local Similarity 99.3%; Pred. No. 0;
XX Matches 2292; Conservative 1; Mismatches 5; Indels 10; Gaps 4;
XX
XX 1 CTGAATGATGCTCCCAAGTCTATGTGATGAGTCCCTTGTGCCCAAGGAAATAGGTG 60
XX
XX 7056 CTGAATGATGCTCCCAAGTCTATGTGATGAGTCCCTTGTGCCCAAGGAAATAGGTG 7115
XX
XX 61 TGTCCATGTCATTTAAAGCATTAATACACTAATATAGTTCTCTCTCTTTGG 120
XX
XX 7116 TGTCCATGTCATTTAAAGCATTAATACACTAATATAGTTCTCTCTCTTTGG 7175
XX
XX 121 ATAATAGGACGTTGTTCCGATGGACTTGAAGCCCTTAGGTCCCAAGAGCCCTTGGAGA 180
XX
XX 7176 ATAATAGGACGTTGTTCCGATGGACTTGAAGCCCTTAGGTCCCAAGAGCCCTTGGAGA 7235
XX
XX 181 ATTACTTCCCACAAAGCACACACCGTGAAATAGATGCCAATGTGCTAGCCAAAAG 240
```

```
Db 7236 ATTTACTTCCCAAGGCAACAGACCGTGAGATAGATGCCAATGTGCTAGCCAAAAG 7295
Qy 241 ACAAGAGTGTCTGGAATTTTTCAGAGCAGGAGAAAGAACTCAGGTGAGCAGAAAACACCTT 300
Db 7296 ACAAGAGTGTCTGGAATTTTTCAGAGCAGGAGAAAGAACTCAGGTGAGCAGAAAACACCTT 7355
Qy 301 TGCCTTTCAATCAGTTTAAACAGCCTCCTGAACTCCTTCTCTATCATGCTACTGCTCTCTG 360
Db 7356 TGCCTTTCAATCAGTTTAAACAGCCTCCTGAACTCCTTCTCTATCATGCTACTGCTCTCTG 7415
Qy 361 TTTTAGAGAGACTAAACAGAGACATTGAAAGTCAGGTAAGCTGAATATAACATTGCTGA 420
Db 7416 TTTTAGAGAGACTAAACAGAGACATTGAAAGTCAGGTAAGCTGAATATAACATTGCTGA 7475
Qy 421 AATGTTTTTCTTGTGTTTAAACAGGGCTCAAGACATTAATGGAGAAAGACTGGAATAA 480
Db 7476 AATGTTTTTCTTGTGTTTAAACAGGGCTCAAGACATTAATGGAGAAAGACTGGAATAA 7535
Qy 481 TCATAAGAAAGGAAAGACTGTTCCAAAGCTTGGGAAAAGTGTATTATCAGCAGTTAGT 540
Db 7536 TCATAAGAAAGGAAAGACTGTTCCAAAGCTTGGGAAAAGTGTATTATCAGCAGTTAGT 7595
Qy 541 GAGAGGAAGAAAATCAGAAAGAGTTTCAGAGGAAACACCTAAGACAAACAGGTAAAGGG 600
Db 7596 GAGAGGAAGAAAATCAGAAAGAGTTTCAGAGGAAACACCTAAGACAAACAGGTAAAGGG 7655
Qy 601 AAGGAAGAAAATTTAGGTAAGAGGTTTCAAGAAACAACTAGCCCGCAGTCAGTGCCAG 660
Db 7656 AAGGAAGAAAATTTAGGTAAGAGGTTTCAAGAAACAACTAGCCCGCAGTCAGTGCCAG 7715
Qy 661 CAGCCTGTTCTCCAGCCCTTCTTACCCTGGGAGGTGAAAGACTTAGAAAACAGTAGCAG 720
Db 7716 CAGCCTGTTCTCCAGCCCTTCTTACCCTGGGAGGTGAAAGACTTAGAAAACAGTAGCAG 7775
Qy 721 AGGAGATCTATCATCTCCTATGATTAAGAGGAGAAAGAAATCCCTCTTAATATTTCCA 780
Db 7776 AGGAGATCTATCATCTCCTATGATTAAGAGGAGAAAGAAATCCCTCTTAATATTTCCA 7835
Qy 781 TGAAGCTCTGGAATCAAAACCGATGCTCTGTACTTTTAGCACAATACATTTCTATCTAC 840
Db 7836 TGAAGCTCTGGAATGCAAAACCGATGCTCTGTACTTTTAGCACAATACATTTCTATCTAC 7895
Qy 841 AGGTAGATTTCCCAACCAAAATATATCCAGAGATGCTTGTCTATGGTGTATATACAGC 900
Db 7896 AGGTAGATTTCCCAACCAAAATATATCCAGAGATGCTTGTCTATGGTGTATATACAGC 7955
Qy 901 CTTTGTCTCTCTGAGTCAATGATTTTACCACTTTCCTCGAGAAATCGAAAATCATTTTGG 960
Db 7956 CTTTGTCTCTCTGAGTCAATGATTTTACCACTTTCCTCGAGAAATCGAAAATCATTTTGG 8015
Qy 961 GGAGCGGACATTTAGAAAAGAAATCAAGTGTCTAGGATTAATCAAAATTTCTCAATAGTT 1020
Db 8016 GGAGCGGACATTTAGAAAAGAAATCAAGTGTCTAGGATTAATCAAAATTTCTCAATAGTT 8075
Qy 1021 GCAGTTATTTAGATGCGCAAGGAAATTAAGTCAATAGATAGGTTGGTAGAATTTAG 1080
Db 8076 GCAGTTATTTAGATGCGCAAGGAAATTAAGTCAATAGATAGGTTGGTAGAATTTAG 8135
Qy 1081 AACATGCTGTTTTTTCAGGTTTATGTC-----TTTTTTTTTTTTTTTTTTTTTAAATAGGG 1135
Db 8136 AACATGCTGTTTTTTCAGGTTTATGTC-----TTTTTTTTTTTTTTTTTTTTTAAATAGGG 8195
Qy 1136 AAATGTTTTTGTGAGAGCCCAATGCTCATTCAAAAGCTCTCTTTTCTGGTTCAGTC 1195
Db 8196 AAATGTTTTTGTGAGAGCCCAATGCTCATTCAAAAGCTCTCTTTTCTGGTTCAGTC 8255
Qy 1196 ATGTGCTGGGACAGAAAGGATCTGGATTAGGCAACATCATAGAGTTGCTCTGAGCTGC 1255
Db 8256 ATGTGCTGGGACAGAAAGGATCTGGATTAGGCAACATCATAGAGTTGCTCTGAGCTGC 8315
Qy 1256 TCTTTGGTGATACCTTCCCAATCTTAACTTTTGGAAATTCACAACTCAAGCTCAAGAGAGA 1315
```


Db 8316 TCTTTGGTGATAACCCCTTCCAAATCCTAAATCTTTTGGAAATTCACAAGCTCAAAAGGGA 8375
Qy 1316 AACCTACTCTGATCTACACATGTTTCGCAATTTTCTATCATGCTCTATGGAACCTC 1375
Db 8376 AACCTACTCTGATCTACACATGTTTCGCAATTTTCTATCATGCTCTATGGAACCTC 8435
Qy 1376 TCTTAGAATCCAGTGGCAGAAGTTCTATGATTAAGAGTTCTGAGCTCAGGCCAGGCA 1435
Db 8436 TCTTAGAATCCAGTGGCAGAAGTTCTATGATTAAGAGTTCTGAGCTCAGGCCAGGCA 8495
Qy 1436 GTCTAGAACTACTCTGAGTTTACTACTGATTTGTTGGGCGAGCTCAGTATCGTT 1495
Db 8496 GTCTAGAACTACTCTGAGTTTACTACTGATTTGTTGGGCGAGCTCAGTATCGTT 8555
Qy 1496 TCTTCACACCTGTTATGAGAGTATCCATATTTATGGTCCGAGGCGCAATGCTCCCA 1555
Db 8556 TCTTCACACCTGTTATGAGAGTATCCATATTTATGGTCCGAGG-CAGTAATGCTCCCA 8614
Qy 1556 CGAGATCAGTTCTGAACCTAACCTGGAATTTTATGAGTTTATGAGTTTATGAGTTTATG 1615
Db 8615 CGAGATCAGTTCTGAACCTAACCTGGAATTTTATGAGTTTATGAGTTTATGAGTTTATG 8674
Qy 1616 AATCAACATTAAGTTCTTCCCTCTGTTTCTCTGTAAGATTAAGCTGCAAAAAA 1675
Db 8675 AATCAACATTAAGTTCTTCCCTCTGTTTCTCTGTAAGATTAAGCTGCAAAAAA 8734
Qy 1676 AAAAAATCTTTTAAAAATGAATGCAATGAATTTGCTCTGGGCTACTGATGCTTC 1735
Db 8735 AAAAAATCTTTTAAAAATGAATGCAATGAATTTGCTCTGGGCTACTGATGCTTC 8794
Qy 1736 TTTTCTTTTCTCTCTTTTCACTAAGTCACTGATGATGATGATGATGATGATGATGATGAT 1795
Db 8795 TTTTCTTTTCTCTCTTTTCACTAAGTCACTGATGATGATGATGATGATGATGATGATGAT 8854
Qy 1796 TCAAAACCTATGCTGAGTTCTCAAGGCGAGGTCATATGATGATGATGATGATGATGATGATG 1855
Db 8855 TCAAAACCTATGCTGAGTTCTCAAGGCGAGGTCATATGATGATGATGATGATGATGATGATG 8914
Qy 1856 CTACGAGAAACCAACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1915
Db 8915 CTACGAGAAACCAACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 8974
Qy 1916 CCTTAGACTTGACATATGACCCCTGCTTCCATTTCTAAGCATAGGCGAGGCTTTAT 1974
Db 8975 CCTTAGACTTGACATATGACCCCTGCTTCCATTTCTAAGCATAGGCGAGGCTTTAT 9034
Qy 1975 TTTACAATGGTAATA---GATATCACTTGAAGTTTATCAAAAGAGTTGCGGCGGTGGTG 2031
Db 9035 TTTACAATGGTAATAGATATCACTTGAAGTTTATCAAAAGAGTTGCGGCGGTGGTG 9094
Qy 2032 AAGATTCAACACAGATTCAGTTTGTGTCAGATTTCTAATTTTACATGTTCTTT 2091
Db 9095 AAGATTCAACACAGATTCAGTTTGTGTCAGATTTCTAATTTTACATGTTCTTT 9154
Qy 2092 TGCCAAAGGGTGATTTTTTAAATAACATTTTCTTTCTTTATCTTTTATGAGTTC 2151
Db 9155 TGCCAAAGGGTGATTTTTTAAATAACATTTTCTTTCTTTATCTTTTATGAGTTC 9214
Qy 2152 GAGACCATGAGAAACAGCGTCAATCTCTTTTCTATGATCCAGCTGAAAGGCAAGCC 2211
Db 9215 GAGACCATGAGAAACAGCGTCAATCTCTTTTCTATGATCCAGCTGAAAGGCAAGCC 9274
Qy 2212 CTCCAGAGAGCGTTATGTGACCCACACAGCGATTTGTTGATGATGATGATGATGATGATG 2271
Db 9275 CTCCAGAGAGCGTTATGTGACCCACACAGCGATTTGTTGATGATGATGATGATGATG 9334
Qy 2272 TCTGAAGCCATAGCTTCCAGGAGGCC 2299
Db 9335 TCTGAAGCCATAGCTTCCAGGAGGCC 9362

RESULT 12
ABV76130

ID ABV76130 standard; cDNA; 1250 BP.
XX AC ABV76130;
DT 07-MAR-2003 (first entry)
XX Human endothelin-1 nucleic acid.
DE DE Gene therapy; vector; hepatitis B virus; cardiovascular disease; heart;
KW KW cardiant; vasotropic; antiarrhythmic; antiarteriosclerotic; human;
KW KW endothelin-1; gene; ss.
XX OS Homo sapiens.
XX PN WO200287594-A1.
XX PD 07-NOV-2002.
XX PF 30-APR-2002; 2002WO-US013644.
XX PR 30-APR-2001; 2001US-0287423P.
XX PA (REGC) UNIV CALIFORNIA.
XX PI Chien KR, Hoshijima M;
XX DR WPI; 2003-111844/10.
XX PT Novel non-viral vector comprises vesicular membrane with hepatitis B
PT envelope protein with cardiac targeting sequence, and nucleotide sequence
PT for gene therapy useful for treating, e.g., heart failure, arrhythmia and
PT atherosclerosis.
XX PS Disclosure; Page 21-22; 53pp; English.
XX CC The present sequence is that of a human endothelin-1 nucleic acid. The
CC invention provides a non-viral vesicle vector for the delivery of nucleic
CC acid to various cardiac cell types. The vesicle vector contains the
CC hepatitis B virus envelope protein in which at least part of the liver
CC targeting sequence is deleted and replaced with a specific cardiac cell
CC targeting sequence, e.g. a vasotropic peptide such as endothelin-1,
CC which is used to broadly target the cardiovascular system. The vesicle
CC vector can be delivered intravenously or intra-arterially rather than by
CC more invasive methods such as direct cardiac injection. It can be used to
CC deliver gene products to replace or enhance expression of proteins for
CC treatment of heart failure, arrhythmia, reperfusion injury,
CC atherosclerosis, to promote angiogenesis, etc. The vesicles are highly
CC stable and can be produced in large quantities, making them ideal for
CC gene therapy
SQ Sequence 1250 BP; 312 A; 332 C; 318 G; 288 T; 0 U; 0 Other;
Query Match 6.98; Score 159; DB 7; Length 1250;
Best Local Similarity 97.0%; Pred. No. 2.7e-27;
Matches 162; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 128 GCACGTTGTTCCGTTATGAGCTTGAAGCCCTAGTCCCAAGAGAGCCCTTGGAGATTACT 187
Db 568 GCACGTTGTTCCGTTATGAGCTTGAAGCCCTAGTCCCAAGAGAGCCCTTGGAGATTACT 627
Qy 188 TCCCAACAAGGCAACAGACCGTGAAATAGATGCAATGTCTAGCCAAAGACACAGAA 247
Db 628 TCCCAACAAGGCAACAGACCGTGAAATAGATGCAATGTCTAGCCAAAGACACAGAA 687
Qy 248 GTGCTGGAATTTTTCACAGCAGGAAAGAACTCAGGTGAGCAGAAA 294
Db 688 GTGCTGGAATTTTTCACAGCAGGAAAGAACTCAGGTGAGCAGAAA 734
RESULT 13
AAH48266
ID AAH48266 standard; DNA; 636 BP.
XX

AC AAH48266;
 XX 21-SEP-2001 (first entry)
 XX Heart muscle cell differentiation related DNA SEQ ID NO: 66.
 DE Heart muscle cell; human; cell differentiation; heart disease; ds.
 KW Heart muscle cell; human; cell differentiation; heart disease; ds.
 XX Homo sapiens.
 OS WO200148151-A1.
 PN 05-JUL-2001.
 PD 27-DEC-2000; 2000WO-JP009323.
 XX 28-DEC-1999; 99JP-00372826.
 PR 28-FEB-2000; 2000WO-JP001148.
 PR 02-NOV-2000; 2000WO-JP007741.
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 PA Umezawa A, Hata J, Fukuda K, Ogawa S, Sakurada K, Gojo S;
 PI Yamada Y;
 XX WPI: 2001-425656/45.
 DR P-PSDB; AAG64862.
 XX Cells capable of differentiating into cardiomyocytes and originating in
 PT bone marrow or umbilical blood cells for study of cardiomyocyte
 PT differentiation and treatment of heart disease.
 XX Claim 32; Page 168-170; 183pp; Japanese.
 XX The present invention provides cells originating in the human bone marrow
 CC or umbilical blood cells which are capable of differentiating into
 CC cardiomyocytes. These cells are useful in the treatment of diseases
 CC involving heart muscle degeneration, such as myocardial infarction, and
 CC the study of cardiomyocyte differentiation. The present sequence is an
 CC oligonucleotide described in the exemplification of the invention
 XX Sequence 636 BP; 191 A; 147 C; 165 G; 133 T; 0 U; 0 Other;
 SQ Query Match 6.8%; Score 157.4; DB 4; Length 636;
 Best Local Similarity 96.4%; Pred. No. 5.1e-27;
 Matches 161; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 128 GCACGTTGTTCCGTATGGACTTGGAGCCCTAGGTCCTCAGAGAGCCCTTGGAGATTACT 187
 Db 234 GCACGTTGTTCCGTATGGACTTGGAGCCCTAGGTCCTCAGAGAGCCCTTGGAGATTACT 293
 QY 188 TCCCAAGGCAACAGACCGCTGAAATAGATGCCAATGTGCTAGCCAAAGACAAGAA 247
 Db 294 TCCCAAGGCAACAGACCGCTGAAATAGATGCCAATGTGCTAGCCAAAGACAAGAA 353
 QY 248 GTGCTGGAATTTTCCCAAGCAGGAAAGAACTAGCTAGGTCAGCAGAA 294
 Db 354 GTGCTGGAATTTTCCCAAGCAGGAAAGAACTAGCTAGGTCAGCAGAA 400
 RESULT 14
 AAH4397
 ID AAH4397 standard; cDNA; 636 BP.
 XX AAH4397;
 XX 26-SEP-2001 (first entry)
 XX Human ET1 nucleotide sequence SEQ ID NO:66.
 DE Differentiation; heart muscle cell; cytokine; transcription factor;
 KW proliferation; surface antigen; heart disease; cardiomyocyte;
 KW bone marrow; umbilical blood cell; heart muscle degeneration;

KW myocardial infarction; ss.
 XX Homo sapiens.
 OS WO200148150-A1.
 PN 05-JUL-2001.
 PD 02-NOV-2000; 2000WO-JP007741.
 XX 28-DEC-1999; 99JP-00372826.
 PR 28-FEB-2000; 2000WO-JP001148.
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 PA Umezawa A, Hata J, Fukuda K, Ogawa S, Sakurada K, Gojo S;
 PI Yamada Y;
 XX WPI: 2001-425655/45.
 DR P-PSDB; AAB99933.
 XX Cells capable of differentiating into cardiomyocytes and originating in
 PT bone marrow or umbilical blood cells for study of cardiomyocyte
 PT differentiation and treatment of heart disease.
 XX Claim 32; Page 162-164; 187pp; Japanese.
 XX The present invention describes cells originating in bone marrow or
 CC umbilical blood cells which are capable of differentiating into
 CC cardiomyocytes. Also described are: (1) cardiomyocytes produced by the
 CC differentiation of the cells; (2) a method for carrying out the
 CC differentiation into cardiomyocytes, regulated by a promotional and/or
 CC inhibitory factor; (3) a method for the differentiation of the cells into
 CC cell types other than cardiomyocytes; (4) drug compositions promoting the
 CC formation of heart muscle and regeneration of heart tissue which contain
 CC the cells; (5) a method for the production of antibodies which recognise
 CC the cells, especially antibodies which recognise a surface antigen on the
 CC cells; (6) a method for screening factors which promote the proliferation
 CC of the cells; (7) a method for immortalising the cells by expressing
 CC telomerase in them; (8) drug compositions for the treatment of heart
 CC disease which contain the immortalised cells; and (9) cell-free
 CC supernatant from the culture of the cells and its use in promoting their
 CC differentiation into cardiomyocytes. The cells are used in the treatment
 CC of diseases involving heart muscle degeneration, such as myocardial
 CC infarction and in the study of cardiomyocyte differentiation. AAH44351 to
 CC AAH44409 and AAB99915 to AAB99935 represent sequences used in the
 CC exemplification of the present invention
 XX Sequence 636 BP; 191 A; 147 C; 165 G; 133 T; 0 U; 0 Other;
 SQ Query Match 6.8%; Score 157.4; DB 4; Length 636;
 Best Local Similarity 96.4%; Pred. No. 5.1e-27;
 Matches 161; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 128 GCACGTTGTTCCGTATGGACTTGGAGCCCTAGGTCCTCAGAGAGCCCTTGGAGATTACT 187
 Db 234 GCACGTTGTTCCGTATGGACTTGGAGCCCTAGGTCCTCAGAGAGCCCTTGGAGATTACT 293
 QY 188 TCCCAAGGCAACAGACCGCTGAAATAGATGCCAATGTGCTAGCCAAAGACAAGAA 247
 Db 294 TCCCAAGGCAACAGACCGCTGAAATAGATGCCAATGTGCTAGCCAAAGACAAGAA 353
 QY 248 GTGCTGGAATTTTCCCAAGCAGGAAAGAACTAGCTAGGTCAGCAGAA 294
 Db 354 GTGCTGGAATTTTCCCAAGCAGGAAAGAACTAGCTAGGTCAGCAGAA 400
 RESULT 15
 AAH35169
 ID AAH35169 standard; DNA; 1167 BP.
 XX AAH35169;
 XX

DT 28-JUL-2000 (first entry)
XX Human adenosine receptor related polynucleotide 2nd SEQ ID NO:43.
DE
XX
XX Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW phosphorothioate; impaired respiration; inflammation; allergy;
KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX
XX Homo sapiens.
OS
XX WO200009525-A2.
PN
XX 24-FEB-2000.
PD
XX
XX 03-AUG-1999; 99WO-US017712.
PF
XX
XX 03-AUG-1998; 98US-0095212P.
PR
XX (UYEC-) UNIV EAST CAROLINA.
PA
XX
XX Nyce JW;
PI
XX WPI; 2000-205971/18.
DR
XX
XX New antisense oligonucleotides useful for treating e.g. pulmonary
PT vasoconstriction, inflammation, allergies, asthma, hypertension,
PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT cancers.
PT
XX
XX Disclosure; Page 1201; 1343pp; English.
PS
XX
XX The present invention describes a new composition comprising an antisense
CC oligonucleotide (ON) with low adenosine (up to 15%), which targets
CC nucleic acids involved in bronchoconstriction, allergies, and/or
CC inflammation. The ON can have antiinflammatory, antiallergic,
CC antiasthmatic, cytostatic and analgesic activities. The compositions are
CC useful for the treatment of diseases associated with inflammation,
CC impaired airways, including lung disease and diseases whose secondary
CC effects afflict the lungs of a subject. They can be used for treating
CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma,
CC impeded respiration, respiratory distress syndrome, pain, cystic
CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
CC carcinomas, and cancers which may metastasise to the lungs, including
CC breast and prostate cancer. The reduction of the adenosine content of the
CC ONs reduces side effects. The A-containing ONs break down with the
CC release of deoxyadenosine which activates adenosine receptors causing
CC bronchoconstriction and inflammation. AAA32913 to AAA35312 represent the
CC nucleotide sequences given in the sequence listing from the present
CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185
CC sequences are also called SEQ ID NO:1 to 185, but the sequences differ
CC from the previously named sequences. SEQ ID NO:11 to 1680 (AAA32323 to
CC AAA33992) are specifically claimed ONs from the present invention. N.B.
CC Sequences given in the disclosure of the present invention do not match
CC up with their corresponding SEQ ID NO: sequences given in the sequence
CC listing
XX
XX
SQ Sequence 1167 BP; 288 A; 309 C; 299 G; 271 T; 0 U; 0 Other;
Query Match 6.8%; Score 157.4; DB 3; Length 1167;
Best Local Similarity 96.4%; Pred. No. 6.3e-27;
Matches 161; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 128 GCACGTTGTTCCCGTATGGACTTGAAGCCCTAGTCCCAAGAGAGCCCTGGAGAAATTACT 187
Db 486 GCACGTTGTTCCCGTATGGACTTGAAGCCCTAGTCCCAAGAGAGCCCTGGAGAAATTACT 545
QY 188 TCCCAACAAGGCAACAGACCGTGAAATAGATGCCAATGTGCTAGCCAAAAGACAGAA 247

Db 546 TCCCAACAAGGCAACAGACCGTGAGAAATAGATGCCAATGTGCTAGCCAAAAGACAGAA 605
QY 248 GTGCTGGAATTTTTCGCAAGCAGGAAAGAACTCAGGTGAGCAGAAA 294
Db 606 GTGCTGGAATTTTTCGCAAGCAGGAAAGAACTCAGGTGAGCAGAAA 652

Search completed: July 16, 2004, 15:57:05
Job time : 875 secs

this Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2004, 15:07:49 ; Search time 162 Seconds
(without alignments)
7882.358 Million cell updates/sec

Title: US-10-043-715-1_COPY_157700_160000

Perfect score: 2301

Sequence: 1 ctgaatgatgtctccaagt.....tagctccaaggagagccct 2301

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 582709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq: *

2: /cgn2_6/ptodata/2/ina/5B.COMB.seq: *

3: /cgn2_6/ptodata/2/ina/6A.COMB.seq: *

4: /cgn2_6/ptodata/2/ina/6B.COMB.seq: *

5: /cgn2_6/ptodata/2/ina/PTUS.COMB.seq: *

6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	66.6	2.9	7218	1 US-08-232-463-14	Sequence 14, Appl
C 2	45.2	2.0	223	4 US-09-621-976-10791	Sequence 10791, A
C 3	45.2	2.0	5152	4 US-10-204-708-48	Sequence 48, Appl
C 4	44.4	1.9	832	4 US-09-621-976-2813	Sequence 2813, Ap
C 5	43.6	1.9	5501	4 US-10-204-708-38	Sequence 38, Appl
C 6	42.8	1.9	7218	1 US-08-232-463-14	Sequence 14, Appl
C 7	42.4	1.8	1492	4 US-09-369-247-23	Sequence 23, Appl
C 8	42	1.8	2668	4 US-09-370-838-156	Sequence 156, App
C 9	41.6	1.8	4597	4 US-08-961-527-175	Sequence 175, App
C 10	40.4	1.8	9347	4 US-10-204-708-35	Sequence 35, Appl
C 11	40.2	1.7	5152	4 US-10-204-708-73	Sequence 73, Appl
C 12	39.6	1.7	535	4 US-09-621-976-2255	Sequence 2255, Ap
C 13	39.6	1.7	5501	4 US-10-204-708-37	Sequence 37, Appl
C 14	39.4	1.7	6306	4 US-10-204-708-50	Sequence 50, Appl
C 15	39	1.7	8093	4 US-10-204-708-32	Sequence 32, Appl
C 16	39	1.7	11049	4 US-10-204-708-23	Sequence 23, Appl
C 17	38.8	1.7	3394	4 US-09-738-946-7	Sequence 7, Appl
C 18	38.6	1.7	586	3 US-09-328-111-366	Sequence 366, App
C 19	38.4	1.7	2148	4 US-09-134-000C-1689	Sequence 1689, Ap
C 20	38.4	1.7	640681	4 US-09-790-988-1	Sequence 1, Appl
C 21	38.2	1.7	474	4 US-09-621-976-18033	Sequence 18033, A
C 22	38.2	1.7	853	3 US-09-174-768-5	Sequence 5, Appl
C 23	38.2	1.7	1353	4 US-08-956-171E-521	Sequence 521, App
C 24	38.2	1.7	786431	4 US-09-751-389-3	Sequence 3, Appl
C 25	38	1.7	1945	3 US-08-714-918-77	Sequence 77, Appl
C 26	38	1.7	1945	3 US-09-265-315-77	Sequence 77, Appl
C 27	38	1.7	1945	3 US-09-265-315-77	Sequence 77, Appl

Query Match 2.9%; Score 66.6; DB 1; Length 7218;

ALIGNMENTS

RESULT 1

US-08-232-463-14/c

; Sequence 14, Application US/08232463

; Patent No. 5670367

; GENERAL INFORMATION:

; APPLICANT: DORNER, F.

; APPLICANT: SCHEIFLINGER, F.

; APPLICANT: FALKNER, F. G.

; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

; NUMBER OF SEQUENCES: 52

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 1800 Diagonal Road, Suite 500

; CITY: Alexandria

; STATE: VA

; COUNTRY: USA

; ZIP: 22313-0299

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA: US/08/232,463

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/935,313

; FILING DATE:

; APPLICATION NUMBER: EP 91 114 300.6

; FILING DATE: 26-AUG-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: BENT, Stephen A.

; REGISTRATION NUMBER: 29,768

; REFERENCE/DOCKET NUMBER: 30472/114 IMM

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703)836-9300

; TELEX: (703)883-4109

; INFORMATION FOR SEQ ID NO: 14:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 7218 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; CLONE: ptz9pt-F1s

; US-08-232-463-14

	Best Local Similarity	5.6%; Pred.No. 4e-07;				
	Matches	18; Conservative	191; Mismatches	110; Indels	0; Gaps	0;
QY	365	AGAGAGACTAACAGAGACATTTGAAGTTCAGGGTAAGCTGAAATATAACATTGCTGAAATG	424			
		: : : : :	: : : : :	: : : : :	: : : : :	
Db	1354	RRR	1295			
QY	425	TTTTTCTTGCTGATTTTAAACAGGGCTGAAGACATTATCGAGAAAGACTCGAATAATCAT	484			
		: : : : :	: : : : :	: : : : :	: : : : :	
Db	1294	RRR	1235			
QY	485	AAGAAAAGGAAAAAGACTGTTCCAAGCTTGGGAAAAAGTAGTATTTATCAGCAGTTAGTGCAGA	544			
		: : : : :	: : : : :	: : : : :	: : : : :	
Db	1234	RRR	1175			
QY	545	GGAAGAAAATCAGAGAAGTTCAGAGGAACA CTAGACAAACCAGGTGAAGAGGAAGG	604			
		: : : : :	: : : : :	: : : : :	: : : : :	
Db	1174	RRR	1115			
QY	605	AAGAAAAATTAGGTAGAGAGTTCCACAAAGAACAACTAGCCCCAGTGTATGCCACAGC	664			
		: : : : :	: : : : :	: : : : :	: : : : :	
Db	1114	RRR	1055			
QY	665	CTGTTCTCTCAGCCCTTCT	683			
		: : : : :	: : : : :	: : : : :	: : : : :	
Db	1054	CTCGACTGCAGCCAAGCT	1036			

```

RESULT 2
US-09-621-976-10791/c
; Sequence 10791, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 10791
; LENGTH: 223
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-10791

```

	Query Match	2.0%	Score 45.2;	DB 4;	Length 223;
	Best Local Similarity 54.1%;		Pred. No. 0.027;		
	Matches 92;	Conservative 0;	Mismatches 78;	Indels 0;	Gaps 0;
QY	1090	TTTTTCAGGTTTATGGTC	TTTTTTTTTTTTTTTTTTTTTTTTTTTT	TTAATAGGGAATGCTTTGGTG	1149
Db	222	TTTTTTTTTTTTTTTTTTTTTTTTTTTT	TTTTTTTTTTTTTTTTTTTTTTTTTTTT	TACGGTAGCAGTTTATTGTC	163
QY	1150	CAGAGCCAAATGTCATTC	CCAAAAAGCTCTCTCTTTTCTGTGTC	ATGTCATGCTCGGACAG	1209
Db	162	TACAAATAATGACAA	TATTAATCTTCCTATGTTCTCTCATCT	GGCTTTGTCAGGGGTT	103
QY	1210	AGAAGGATCTGGATTG	AGGCAACATATAGAGTTGCTCT	GAGCTGCTCTTT	1259
Db	102	AGTAGTCCTGATTTTG	TATATTTAAAAAAATAGTCATCGGTCG	TCCTT	53

RESULT 3
US-AO-204-708-48
; Source 48, Application US/10204708
; Security No. 667731
; Reference No. 667731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt

[illegible]

```

RESULT 4
US-09-621-976-2813
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399.
US-09-621-976-2813
Query Match      1.9%  Score 44.4;  DB 4;  Length 832;
Best Local Similarity 14.1%;  Pred. No. 0.077;
Matches 54;  Conservative 156;  Mismatches 172;  Indels 0;  Gaps 0;
992 TCATGGATAATCAAATCTTCATATAAGTTCGAGTTATTCAGATGGCCAAAGGAAAAATAA 1051

```



```

Db      3484 AAAAAAAAAAAAAAAAAAATTAAGTTTTTTTATGTTATAAGTTTTATGTGA   3543
QY      1095 CAGGTTTATGGCTTTTTTTTTTTTTTTTTTTTTTTTAAATA   1132
Db      3544 TTTATTTTTTGTTGTTTTTTTTTATTGTTATTATTGTTA   3581

RESULT 6
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:

```

```

> OPERATING SYSTEM: PC-DOS/MS-DOS
> SOFTWARE: PatentIn Release #1.0, Version #1.25
> CURRENT APPLICATION DATA: US/08/232,463
> APPLICATION NUMBER: US/08/232,463
> FILING DATE:
> CLASSIFICATION: 435
> PRIOR APPLICATION DATA:
> APPLICATION NUMBER: US/07/935,313
> FILING DATE:
> APPLICATION NUMBER: EP 91 114 300.6
> FILING DATE: 26-AUG-1991
> ATTORNEY/AGENT INFORMATION:
> NAME: BENT, Stephen A.
> REGISTRATION NUMBER: 29,768
> REFERENCE/DOCKET NUMBER: 30472/114 IMMU
> TELECOMMUNICATION INFORMATION:
> TELEPHONE: (703)836-9300
> TELEFAX: (703)683-4109
> TELEX: 899149
> INFORMATION FOR SEQ ID NO: 14:
> SEQUENCE CHARACTERISTICS:
> LENGTH: 7218 base pairs
> TYPE: nucleic acid
> STRANDEDNESS: single
> TOPOLOGY: linear
> IMMEDIATE SOURCE:
> CLONE: pTZgpt-Fis
> US-08-232-463-14

```

[illegible]

[illegible]

```

RESULT 7
US-09-369-247-23/c
; Sequence 23, Application US/09369247
; Patent No. 6569992
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 44 Human Secreted Proteins
; FILE REFERENCE: PZ024P1
; CURRENT APPLICATION NUMBER: US/09/369,247
; CURRENT FILING DATE: 1999-08-05
; EARLIER APPLICATION NUMBER: 60/074,118
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/074,157
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/074,137
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/074,341
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/074,141
; EARLIER FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 1492
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-369-247-23

```

	Query Match	1.8%;	Score 42.4;	DB 4;	Length 1492;
	Best Local Similarity	56.7%;	Pred. No. 0.32;		
	Matches 76;	Conservative 1;	Mismatches 57;	Indels 0;	Gaps 0;
Qy	1088	TGTTTTTCAGGTTTATGGTCTTTTTTTTTTTTTTTTTTTTTTAAATAGGGAAATGCTGTTTCG	1147		
Ddb	1474	TTT	1415		
Qy	1148	TGCAGAGCCAAATGTCATTCAAAAAGCTCTCTTTTCTCTGGTCAGTCATGTGCTGGGAC	1207		
Ddb	1414	TTGAAACCAATAATTATCAAAACGCTGGTGTGTATGTGGGGGGAGGGTGTCCGAC	1355		
Qy	1208	AGAGAAAGGATCTG	1221		
Ddb	1354	AGACAGGGCAGCGG	1341		

RESULT 8
US-09-370-838-156/c
; Sequence 156, Application US/09370838
; Patent No. 6444425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.

```

/ APPLICANT: Mohamath, Roadoh
/
/ APPLICANT: Secrist, Heather
/
/ TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
/
/ TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
/
/ FILE REFERENCE: 210121.475C1
/
/ CURRENT APPLICATION NUMBER: US/09/370,838
/
/ CURRENT FILING DATE: 1999-08-09
/
/ EARLIER APPLICATION NUMBER: US 09/285,323
/
/ EARLIER FILING DATE: 1999-04-02
/
/ NUMBER OF SEQ ID NOS: 289
/
/ SOFTWARE: FastSEQ for Windows Version 3.0
/
/ SEQ ID NO 156
/
/ LENGTH: 2668
/
/ TYPE: DNA
/
/ ORGANISM: Homo sapien
/
/ US-09-370-838-156

```

	Query Match	1.8%	Score 42;	DB 4;	Length 2668;
	Best Local Similarity	56.7%;	Pred. NO. 0.52;		
	Matches 60;	Conservative 0;	Mismatches 30;	Indels 0;	Gaps 0;
QY	1071	TAGAATTAGAACATGCTGTTTTTCAGGTTATGCTCTTTTTTTTTTTTTTTTTTTAA	1130		
DB	2667	TCGAGTATTTGA	2608		
QY	1131	TAGGAAATGTGTTGGTCAGAGCCAATG	1160		
DB	2607	AAATTAATGATTTATTGCAGGGCCAATG	2578		

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match
1.8%; Score 41.6; DB 4; Length 4597;

Search completed: July 16, 2004, 17:36:50
Job time : 165 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2004, 17:34:04 ; Search time 1001 Seconds
(without alignments)
11214.498 Million cell updates/sec

Title: US-10-043-715-1_COPY_157700_160000
Perfect score: 2301
Sequence: 1 ctgaaatgagctcccaagt.....tagcctccacgagagccct 2301

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3190992 seqs, 2439311697 residues

Total number of hits satisfying chosen parameters: 6381994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/PCTUS_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	ID	Description
1	2301	100.0	186510	15	US-10-043-715-1
2	727.4	31.6	1339	13	US-10-027-632-124859
3	727.4	31.6	1339	16	US-10-027-632-124859
4	159	6.9	1250	15	US-10-136-819-3
5	157.4	6.8	636	9	US-09-749-728B-66
6	157.4	6.8	1166	15	US-10-295-942-1
7	157.4	6.8	1251	13	US-10-211-462-126
8	157.4	6.8	1251	13	US-10-342-887-701
9	157.4	6.8	1251	13	US-10-172-118-701
10	157.4	6.8	1251	13	US-10-170-385-398
11	157.4	6.8	1251	15	US-10-007-926A-458
12	157.4	6.8	1251	15	US-10-101-510-5
13	157.4	6.8	1251	15	US-10-021-660-40
14	156.4	6.8	566	15	US-10-060-036-22

15	155.8	6.8	512	10	US-09-918-995-21013	Sequence 21013, A
16	143	6.2	266	9	US-09-923-876-5097	Sequence 5097, Ap
17	143	6.2	266	11	US-09-923-876-5097	Sequence 5097, Ap
18	97	4.2	300	9	US-09-294-093B-2321	Sequence 2321, Ap
19	60	2.6	60	10	US-09-908-975-14247	Sequence 14247, A
20	48.6	2.1	6247	17	US-10-257-1665-4	Sequence 4, Appli
21	48.6	2.1	11371	15	US-10-311-455-746	Sequence 746, App
22	47.4	2.1	16287	13	US-10-311-455-645	Sequence 645, App
23	47	2.0	885	13	US-10-027-632-35082	Sequence 35082, A
24	47	2.0	885	16	US-10-027-632-35082	Sequence 35082, A
25	46	2.0	5397	15	US-10-311-455-1018	Sequence 1018, Ap
26	45.6	2.0	14542	13	US-10-311-455-1018	Sequence 105, App
27	45.4	2.0	6076	13	US-10-221-613-105	Sequence 386, App
28	45.4	2.0	10713	15	US-10-221-714A-386	Sequence 714, App
29	45.4	2.0	3673778	15	US-10-311-455-714	Sequence 2, Appli
30	45.2	2.0	5152	15	US-10-312-841-2	Sequence 48, Appli
31	45.2	2.0	13574	15	US-10-204-708-48	Sequence 1289, Ap
32	45	2.0	8237	13	US-10-311-455-1289	Sequence 527, App
33	44.8	1.9	42061	17	US-10-221-714A-527	Sequence 770, App
34	44.6	1.9	11334	15	US-10-322-281-770	Sequence 1205, Ap
35	44.4	1.9	516	17	US-10-311-455-1205	Sequence 6457, Ap
36	44.2	1.9	17527	13	US-10-021-323-6457	Sequence 28, Appli
37	44.2	1.9	17527	15	US-10-240-454-28	Sequence 1406, Ap
38	44.2	1.9	3673778	15	US-10-311-455-1406	Sequence 1, Appli
39	44	1.9	6381	15	US-10-312-841-1	Sequence 679, App
40	43.8	1.9	960	15	US-10-198-846-6381	Sequence 187249, Ap
41	43.6	1.9	476	13	US-10-027-632-187249	Sequence 187250, Ap
42	43.6	1.9	476	13	US-10-027-632-187250	Sequence 187249, Ap
43	43.6	1.9	476	16	US-10-027-632-187249	Sequence 187250, Ap
44	43.6	1.9	476	16	US-10-027-632-187250	Sequence 14764, A
45	43.6	1.9	526	17	US-10-021-323-14764	

ALIGNMENTS

RESULT 1

US-10-043-715-1
; Sequence 1, Application US/10043715
; Publication No. US20030143544A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Jeanette
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF VASCULAR
; TITLE OF INVENTION: DISEASE
; FILE REFERENCE: MMI-009
; CURRENT APPLICATION NUMBER: US/10/043,715
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 186510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-043-715-1

Query Match	100.0%;	Score 2301;	DB 15;	Length 186510;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2301;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	CTGAAATGATGCTCCCAAGTCTATGTGATGAGCTCTTGTGTCCTCCAGTGGATAGGTG	60	
Db	157700	CTGAAATGATGCTCCCAAGTCTATGTGATGAGCTCTTGTGTCCTCCAGTGGATAGGTG	157759	
QY	61	TGTCCTGCTGCATTTTAAAGACTATTAACTAATATAGTTTCTCTCTCTTTGG	120	
Db	157760	TGTCCTGCTGCATTTTAAAGACTATTAACTAATATAGTTTCTCTCTCTTTGG	157819	
QY	121	ATAATAGGACGTTGTTTCCGATGAGCTTGAAGCCCTAGGTCCAAAGAGCCCTTGAGA	180	
Db	157820	ATAATAGGACGTTGTTTCCGATGAGCTTGAAGCCCTAGGTCCAAAGAGCCCTTGAGA	157879	
QY	181	ATTACTTCCCAAGGACACACCGTGAAATAGATGCTAGTCCCAAAAG	240	

157880	DB	ATTTACTTCCCAACAAAGGCCAACAGACCGTGAAAAATAGATGCTAGCCAAAAG	157939
241	QY	ACAAGAAGTGCTCGAAATTTTTTGCACAGCAGGAAAAAGAACTCAGGTGAGCAGAACACCTT	300
157940	DB	ACAAGAAGTGCTCGAAATTTTTTGCACAGCAGGAAAAAGAACTCAGGTGAGCAGAACACCTT	157999
301	QY	TGCTTTTCAATCAGTTTAAACGCCCTCTGAACTCCTTCTCATCATGGTACTGCTCTTCTG	360
158000	DB	TGCTTTTCAATCAGTTTAAACGCCCTCTGAACTCCTTCTCATCATGGTACTGCTCTTCTG	158059
361	QY	TTTTTAGAGAGACTAAACAGAGACATTCAAAGTTCAGGTTAAAGCTCAATATACATTCGCTGA	420
158060	DB	TTTTTAGAGAGACTAAACAGAGACATTCAAAGTTCAGGTTAAAGCTCAATATACATTCGCTGA	158119
421	QY	AATGTTTTTCTTGTGTATTTTAAACAGGGCTGAAGACATTTATGGAGAAAGACTGGAATAA	480
158120	DB	AATGTTTTTCTTGTGTATTTTAAACAGGGCTGAAGACATTTATGGAGAAAGACTGGAATAA	158179
481	QY	TCATAGAAGGAAAGACATGTTCCAAGCTTGGGAAAAAGTGTATTTATCAGCAGTTAGT	540
158180	DB	TCATAGAAGGAAAGACATGTTCCAAGCTTGGGAAAAAGTGTATTTATCAGCAGTTAGT	158239
541	QY	GAGAGGAAGAAAAATCAGAAAGTTCAGAGGAAACACCTTAAGACAAACACAGGTAAAGGG	600
158240	DB	GAGAGGAAGAAAAATCAGAAAGTTCAGAGGAAACACCTTAAGACAAACACAGGTAAAGGG	158299
601	QY	AAGCAAGAAAAATTAGTTAAGAGTTTCACAGACAACTAGTCCCGCTCAGTCATGCCAG	660
158300	DB	AAGCAAGAAAAATTAGTTAAGAGTTTCACAGACAACTAGTCCCGCTCAGTCATGCCAG	158359
661	QY	CAGCCTGTTCTTCAGCCCTTCTTACC CGGCGAGTGAAGACTTGAAGAAACAGTACAG	720
158360	DB	CAGCCTGTTCTTCAGCCCTTCTTACC CGGCGAGTGAAGACTTGAAGAAACAGTACAG	158419
721	QY	AGGAGATCTATGCATCTATAGATTAAGAGGACAAAGAAATCCCTCTTAAATATTTCCA	780
158420	DB	AGGAGATCTATGCATCTATAGATTAAGAGGACAAAGAAATCCCTCTTAAATATTTCCA	158479
781	QY	TGAGCTCTGGAAATCGAAACCGATGCTCTCTGTACTTTTAGCACATACCATTTCATCTAC	840
158480	DB	TGAGCTCTGGAAATCGAAACCGATGCTCTCTGTACTTTTAGCACATACCATTTCATCTAC	158539
841	QY	AGGTAGATTTCCCAACCAAAATATATCAGAGATGTCCTTTGTCAATGGGTATATACAGC	900
158540	DB	AGGTAGATTTCCCAACCAAAATATATCAGAGATGTCCTTTGTCAATGGGTATATACAGC	158599
901	QY	CTTTGCTCTCTGAGTCATGTATTTACCACTTTCCCTGAGAAATCGAAATCATTTTGG	960
158600	DB	CTTTGCTCTCTGAGTCATGTATTTACCACTTTCCCTGAGAAATCGAAATCATTTTGG	158659
961	QY	GGAGCGGACATTTAGAAAAAGAAATCAAAGTGCATGGATTAATCAAATCTTCAATAAGTT	1020
158660	DB	GGAGCGGACATTTAGAAAAAGAAATCAAAGTGCATGGATTAATCAAATCTTCAATAAGTT	158719
1021	QY	GCAGTTATTCAGATGCCAAAGGAAAAATAAGATCATTAGTAGGGTGGTGAATTTAG	1080
158720	DB	GCAGTTATTCAGATGCCAAAGGAAAAATAAGATCATTAGTAGGGTGGTGAATTTAG	158779
1081	QY	AACATGCTGTTTTTTCAGGTTTATGGTCTTTTTTTTTTTTTTTTTTTTTTAAATAGGGAATG	1140
158780	DB	AACATGCTGTTTTTTCAGGTTTATGGTCTTTTTTTTTTTTTTTTTTTTTTAAATAGGGAATG	158839
1141	QY	TGTTTTGGTGCAGAGCCAATGTCAATTCOAAAAGCTCTCTCTTTTCTGCTCAGTCATGTG	1200
158840	DB	TGTTTTGGTGCAGAGCCAATGTCAATTCOAAAAGCTCTCTCTTTTCTGCTCAGTCATGTG	158899
1201	QY	CTGGGACAGAGAGGGATCTCGAATTAGCGAACATCATAGAGTTGCTCTGAGCTGCTCTTT	1260
158900	DB	CTGGGACAGAGAGGGATCTCGAATTAGCGAACATCATAGAGTTGCTCTGAGCTGCTCTTT	158959
1261	QY	GGGTATATACCTTCCAAATCTTAACTTTTTTGGAAATCAACAGCTCAAGAGGAGGAACCT	1320
158960	DB	GGGTATATACCTTCCAAATCTTAACTTTTTTGGAAATCAACAGCTCAAGAGGAGGAACCT	159019


```
Db 911 TGCCTTTCAATCACTGTTTAAACAGCCTCCTCGAACTCCTCTCCATCATGGTACGCTTCCTG 970
QY 361 TTTTAGAGAGACTAAACAGAGACATTGAAAGTCAGGTAAGCTGAATATATACATTGCTGA 420
Db 971 TTTTAGAGAGACTAAACAGAGACATTGAAAGTCAGGTAAGCTGAATATATACATTGCTGA 1030
QY 421 AATGTTTTTCCCTGTTGTTATTTTAAACAGGCTGAAACATATATGGAAGAAAGCTGAATAA 480
Db 1031 AATGTTTTTCCCTGTTGTTATTTTAAACAGGCTGAAACATATATGGAAGAAAGCTGAATAA 1090
QY 481 TCATAAGAAAGAAAGAAAGCTGTTTCAAGCTTCGGGAAAAGTGTATTTATCAGCAGTTAGT 540
Db 1091 TCATAAGAAAGAAAGAAAGCTGTTTCAAGCTTCGGGAAAAGTGTATTTATCAGCAGTTAGT 1150
QY 541 GAGAGAGAGAAATTAACAGAAAGTTTCAAGAAACACCTTAAGACAAACACAGGTAAGAGGG 600
Db 1151 GAGAGAGAGAAATTAACAGAAAGTTTCAAGAAACACCTTAAGACAAACACAGGTAAGAGGG 1210
QY 601 AAGGAAGAAATTAAGTTAAGGTTTCAAGAAACACCTTAAGACAAACACAGGTAAGAGGG 660
Db 1211 AAGGAAGAAATTAAGTTAAGGTTTCAAGAAACACCTTAAGACAAACACAGGTAAGAGGG 1270
QY 661 CAGCCTGTTTCCCTCAGCCCTTCTTACCCGGGAGGTGAAAGACTTAGAAACACAGTAGCAG 720
Db 1271 CAGCCTGTTTCCCTCAGCCCTTCTTACCCGGGAGGTGAAAGACTTAGAAACACAGTAGCAG 1330
QY 721 AGGAGATCT 729
Db 1331 AGGAGATCT 1339

RESULT 4
US-10-136-819-3
; Sequence 3, Application US/10136819
; Publication No. US20030166593A1
; GENERAL INFORMATION:
; APPLICANT: Chien, Kenneth
; TITLE OF INVENTION: No. US20030166593A1-viral vesicle vector for cardiac specific gen
; FILE REFERENCE: 6627-PAL198
; CURRENT APPLICATION NUMBER: US/10/136,819
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: 60/287,423
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1250
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-136-819-3
Query Match 6.9%; Score 159; DB 15; Length 1250;
Best Local Similarity 97.0%; Pred. No. 5.3e-31;
Matches 162; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 128 GCAGTGTTCCTGATGAGCTTGAAGCCCTAGTCCCAAGAGAGCCCTTGAGAAATTTACT 187
Db 568 GCAGTGTTCCTGATGAGCTTGAAGCCCTAGTCCCAAGAGAGCCCTTGAGAAATTTACT 627
QY 188 TCCCAAGAGCAACAGACCCGTGAAATAGATGCAATGTGCTAGCCCAAGAAAGCAAGAA 247
Db 628 TCCCAAGAGCAACAGACCCGTGAAATAGATGCAATGTGCTAGCCCAAGAAAGCAAGAA 687
QY 248 GTGCTGGAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAGAA 294
Db 688 GTGCTGGAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAGAA 734

RESULT 5
US-09-749-728B-66
; Sequence 66, Application US/09749728B
; Patent No. US20020142457A1
; Query Match 6.8%; Score 157.4; DB 15; Length 1166;
; Best Local Similarity 96.4%; Pred. No. 1.3e-30;
; Matches 161; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

; GENERAL INFORMATION:
; APPLICANT: Umezawa, Akihiro
; APPLICANT: Hata, Jun-ichi
; APPLICANT: Fukuda, Keiichi
; APPLICANT: Ogawa, Satoshi
; APPLICANT: Sakurada, Kazuhiro
; APPLICANT: Gojo, Satoshi
; APPLICANT: Yamada, Yoji
; TITLE OF INVENTION: THE CELL HAVING THE POTENTIALITY OF DIFFERENTIATION INTO CARDIOMY
; FILE REFERENCE: 00766.000043
; CURRENT APPLICATION NUMBER: US/09/749,728B
; CURRENT FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: H11-372826
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: PCT-JP00-01148
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: PCT-JP00-07741
; PRIOR FILING DATE: 2000-11-02
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver.2.0
; SEQ ID NO 66
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; OTHER INFORMATION: (1)..(639)
US-09-749-728B-66
Query Match 6.8%; Score 157.4; DB 9; Length 636;
Best Local Similarity 96.4%; Pred. No. 8.8e-31;
Matches 161; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 128 GCAGTGTTCCTGATGAGCTTGAAGCCCTAGTCCCAAGAGAGCCCTTGAGAAATTTACT 187
Db 234 GCAGTGTTCCTGATGAGCTTGAAGCCCTAGTCCCAAGAGAGCCCTTGAGAAATTTACT 293
QY 188 TCCCAAGAGCAACAGACCCGTGAAATAGATGCAATGTGCTAGCCCAAGAAAGCAAGAA 247
Db 294 TCCCAAGAGCAACAGACCCGTGAAATAGATGCAATGTGCTAGCCCAAGAAAGCAAGAA 353
QY 248 GTGCTGGAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAGAA 294
Db 354 GTGCTGGAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAGAA 400

RESULT 6
US-10-295-942-1
; Sequence 1, Application US/10295942
; Publication No. US20030109480A1
; GENERAL INFORMATION:
; APPLICANT: Corder, Roger
; APPLICANT: Smith, Adrian
; APPLICANT: Higenbottom, Tim
; APPLICANT: Rothblatt, Martine
; APPLICANT: Vane, John
; APPLICANT: Jones, Delphine
; TITLE OF INVENTION: INHIBITORS OF ENDOTHELIN-1 SYNTHESIS
; FILE REFERENCE: 080618/0123
; CURRENT APPLICATION NUMBER: US/10/295,942
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US/09/527,240
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1166
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-942-1
Query Match 6.8%; Score 157.4; DB 15; Length 1166;
Best Local Similarity 96.4%; Pred. No. 1.3e-30;
; Matches 161; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

Matches 161; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 128 GCACGTTGTTCCGATGAGCTTGAAGCCCTAGTCCCAAGAGAGCCCTTGAGAAATTACT 187
Db 485 GCACGTTGTTCCGATGAGCTTGAAGCCCTAGTCCCAAGAGAGCCCTTGAGAAATTACT 544
Qy 188 TCCCAAAAGGCAACAGCCGTTGAAATAGATGCAATGCTAGTCCCAAGAGAGCCCTTGAGAAATTACT 247
Db 545 TCCCAAAAGGCAACAGCCGTTGAAATAGATGCAATGCTAGTCCCAAGAGAGCCCTTGAGAAATTACT 604
Qy 248 GTGCTGGAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAGAA 294
Db 605 GTGCTGGAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAGAA 651

RESULT 7
US-10-211-462-126
; Sequence 126, Application US/10211462
; Publication No. US20040033495A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
; TITLE OF INVENTION: Methods of Screening for Angiogenesis Modulators
; FILE REFERENCE: 018501-006200US
; CURRENT APPLICATION NUMBER: US/10/211,462
; PRIOR FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US 09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/791,390
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/310,025
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/334,244
; PRIOR FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 126
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-211-462-126

Query Match 6.8%; Score 157.4; DB 13; Length 1251;
Best Local Similarity 96.4%; Pred. No. 1.4e-30;
Matches 161; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 128 GCACGTTGTTCCGATGAGCTTGAAGCCCTAGTCCCAAGAGAGCCCTTGAGAAATTACT 187
Db 570 GCACGTTGTTCCGATGAGCTTGAAGCCCTAGTCCCAAGAGAGCCCTTGAGAAATTACT 629
Qy 188 TCCCAAAAGGCAACAGCCGTTGAAATAGATGCAATGCTAGTCCCAAGAGAGCCCTTGAGAAATTACT 247
Db 630 TCCCAAAAGGCAACAGCCGTTGAAATAGATGCAATGCTAGTCCCAAGAGAGCCCTTGAGAAATTACT 689
Qy 248 GTGCTGGAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAGAA 294
Db 690 GTGCTGGAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAGAA 736

RESULT 8
US-10-342-887-701
; Sequence 701, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.

; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 701
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-701

Query Match 6.8%; Score 157.4; DB 13; Length 1251;
Best Local Similarity 96.4%; Pred. No. 1.4e-30;
Matches 161; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 128 GCACGTTGTTCCGATGAGCTTGAAGCCCTAGTCCCAAGAGAGCCCTTGAGAAATTACT 187
Db 570 GCACGTTGTTCCGATGAGCTTGAAGCCCTAGTCCCAAGAGAGCCCTTGAGAAATTACT 629
Qy 188 TCCCAAAAGGCAACAGCCGTTGAAATAGATGCAATGCTAGTCCCAAGAGAGCCCTTGAGAAATTACT 247
Db 630 TCCCAAAAGGCAACAGCCGTTGAAATAGATGCAATGCTAGTCCCAAGAGAGCCCTTGAGAAATTACT 689
Qy 248 GTGCTGGAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAGAA 294
Db 690 GTGCTGGAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAGAA 736

RESULT 9
US-10-172-118-701
; Sequence 701, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-178-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 701
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_001955
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-701

Query Match 6.8%; Score 157.4; DB 13; Length 1251;
Best Local Similarity 96.4%; Pred. No. 1.4e-30;
Matches 161; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 128 GCACGTTGTTCCGATGAGCTTGAAGCCCTAGTCCCAAGAGAGCCCTTGAGAAATTACT 187
Db 570 GCACGTTGTTCCGATGAGCTTGAAGCCCTAGTCCCAAGAGAGCCCTTGAGAAATTACT 629

Qy	188	TCCCAAAAGGCAACAGACCGGTGAAAAATAGATGTGCAATGTGCTAGCCAAAAGACCAAGAA	247
Db	630	TCCCAAAAGGCAACAGACCGGTGAAATAGATGTGCAATGTGCTAGCCAAAAGACCAAGAA	689
Qy	248	GTGCTGGAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAGAA	294
Db	690	GTGCTGGAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGACAA	736

```

RESULT 10
US-10-170-385-398
; Sequence 398, Application US/10170385
; Publication No. US20030203372A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Neil Raymond
; APPLICANT: Mundy, Christopher Robert
; APPLICANT: Kan, On
; APPLICANT: Harris, Robert Alan
; APPLICANT: White, Jonathan
; APPLICANT: Binley, Katie Mary
; APPLICANT: Rayner, William Nigel
; APPLICANT: Naylor, Stuart
; APPLICANT: Kingsman, Susan Mary
; APPLICANT: Krige, David
; TITLE OF INVENTION: ANALYSIS METHOD
; FILE REFERENCE: 53268200100
; CURRENT APPLICATION NUMBER: US/10/170,385
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: PCT/GB02/01662
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/05458
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 549
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 398
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-170-385-398

```

Query Match	6.8%;	Score 157.4;	DB 13;	Length 1251;
Best Local Similarity	96.4%;	Pred. No. 1.4e-30;		
Matches 161;	Conservative 0;	Mismatches 6;	Indels 0;	Gaps 0;
OY	128	GCAGGTGTTCCGTTATGGACTTGGAGCCCTAGGTC	CAAGAGAGCGCTTGGAGAA	TTTACT 187
dbb	570	GCAGGTGTTCCGTTATGGACTTGGAGCCCTAGGTC	CAAGAGAGCGCTTGGAGAA	TTTACT 629
OY	188	TCCCACAAAGGCAACAGACCGTGAAATAGATGCCAATGTGTAGC	CAAAAAGACAGAA	247
dbb	630	TCCCACAAAGGCAACAGACCGTGAGATAGATGCCAATGTGTAGC	CAAAAAGACAGAA	689
OY	248	GTGCTGGAATTTTTCGACAGAGAAAAGAACTCAGGTCAGCAGAA		294
db	690	GTGCTGGAATTTTTCGACAGAGAAAAGAACTCAGGCTGAAGACA		736

RESULT 11
 US-10-007-926A-458
 ; Sequence 458, Application US/10007926A
 ; Publication No. US20030143539A1
 ; GENERAL INFORMATION:
 ; APPLICANT: BERTUCCI, FRANCOIS
 ; APPLICANT: HOULGATTE, REMI
 ; APPLICANT: BIRNBAUM, DANIEL
 ; APPLICANT: NGUYEN, CATHERINE
 ; APPLICANT: VIENS, PATRICE
 ; APPLICANT: FERT, VINCENT
 TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
 TITLE OF INVENTION: USING ARRAYS OF CANDIDATE GENES
 FILE REFERENCE: 1546-R-00
 CURRENT APPLICATION NUMBER: US/10/007,926A

```

; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/254,090
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 468
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 458
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 3' terminal sequence
; FEATURE:
; OTHER INFORMATION: endothelin 1 (EDN1)
US-10-007-926A-458

```

	Query Match	6.8%	Score 157.4;	DB 15;	Length 1251;
	Best Local Similarity	96.4%;	Pred. No. 1.4e-30;		
	Matches 161;	Conservative 0;	Mismatches 6;	Indels 0;	Gaps 0;
QY	128	CGAGTGTTC	CGGTATGGACTTTGGAGCCCTAGGTC	CCAGAGAGCGCTTGGAGAA	TTTACT 187
DB	570	GCAGCTGTTC	CGGTATGGACTTTGGAGCCCTAGGTC	CCAGAGAGCGCTTGGAGAA	TTTACT 629
QY	188	TCCACAAAG	CGCAACAGACCGTGA	AAATAGATGCCAATGTGTAGC	CAAAAGACAGAA 247
DB	630	TCCACAAAG	CGCAACAGACCGTGC	GAATAGATGCCAATGTGTAGC	CAAAAGACAGAA 689
QY	248	GTGCTGAA	TTTTTTC	CAACAGCAAGAAAGAACTC	CAGGTGAGCAGAA 294
DB	690	GTGCTGAA	TTTTTTC	CAACAGCAAGAAAGAACTC	CAGGTGAGCAGAA 736

```

RESULT 12
US-10-101-510-5
; Sequence 5, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WAN, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-5

```

Query Match	6.8%;	Score 157.4;	DB 15;	Length 1251;
Best Local Similarity	96.4%;	Prod. No. 1.4e-30;		
Matches 161;	Conservative	0; Mismatch 6;	Indels 0;	Gaps 0;
128 GCAGTGTTCCTATGGACTTGGAGGCCCTAGGTCGAAGAGAGCCCTGGGAATTTTACT	187			
570 GCAGTGTTCCTATGGACTTGGAGGCCCTAGGTCGAAGAGAGCCCTGGGAATTTTACT	629			
188 TCCCAACAAAGGCAACAGACCGTGAAAAATAGATGCCAATGTCTAGCCAAAGAACACAGAA	247			
630 TCCCAACAAAGGCAACAGACCGTGGAATAGATGCCAATGTCTAGCCAAAGAACACAGAA	689			
248 GTGCTGGAAATTTTGCACAGACAGAAAAAGAACTCAGGTGACGAGAA	294			
690 GTGCTGGAAATTTTGCACAGACAGAAAAAGAACTCAGGGCTGAACACA	736			

RESULT 13
US-10-021-650-40
; Sequence 40, Application US/10021660


```
Publication No. US20030152926A1
GENERAL INFORMATION:
APPLICANT: Murray, Richard
APPLICANT: Glynn, Richard
APPLICANT: Watson, Susan R.
APPLICANT: EOS Biotechnology, Inc.
TITLE OF INVENTION: No. US20030152926A1 Methods of Diagnosis of Angiogenesis,
Compositions and Methods of Screening for Angiogenesis
TITLE OF INVENTION: Modulators
FILE REFERENCE: 019501-000710US
CURRENT APPLICATION NUMBER: US/10/021.660
CURRENT FILING DATE: 2001-12-06
PRIOR APPLICATION NUMBER: US/09/784.356
PRIOR FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: US 09/637,977
PRIOR FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 135
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 40
LENGTH: 1251
TYPE: DNA
ORGANISM: Homo sapiens
US-10-021-660-40

Query Match 6.8%; Score 157.4; DB 15; Length 1251;
Best Local Similarity 96.4%; Pred. No. 1.4e-30;
Matches 161; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 128 GCACGTTGTTCCGTATGACCTTGGAAAGCCCTAGTCCCAAGAGAGCCCTTGGAGAAATTTACT 187
DB 570 GCACGTTGTTCCGTATGACCTTGGAAAGCCCTAGTCCCAAGAGAGCCCTTGGAGAAATTTACT 629

QY 188 TCCCAAGGCAACAGACCGTGAAATAGATGCCAATGTCTAGCCCAAGAAAGACAAGAA 247
DB 630 TCCCAAGGCAACAGACCGTGAGATAGATGCCAATGTCTAGCCCAAGAAAGACAAGAA 689

QY 248 GTGCTGGAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAGAAA 294
DB 690 GTGCTGGAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAGAAA 736

RESULT 14
US-10-060-036-22
Publication No. US/10060036
GENERAL INFORMATION:
APPLICANT: Benson, Darin R.
APPLICANT: Kalos, Michael D.
APPLICANT: Lodes, Michael J.
APPLICANT: Persing, David H.
APPLICANT: Hepler, William T.
APPLICANT: Jiang, Yugu
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
AND DIAGNOSIS OF PANCREATIC CANCER
FILE REFERENCE: 210121.566
CURRENT APPLICATION NUMBER: US/10/060.036
CURRENT FILING DATE: 2002-01-30
NUMBER OF SEQ ID NOS: 4560
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 22
LENGTH: 566
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: misc feature
LOCATION: 508, 553
OTHER INFORMATION: n = A,T,C or G
US-10-060-036-22

Query Match 6.8%; Score 156.4; DB 15; Length 566;
Best Local Similarity 96.4%; Pred. No. 1.5e-30;
Matches 160; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 129 CACGTTGTTCCGTATGACCTTGGAAAGCCCTAGTCCCAAGAGAGCCCTTGGAGAAATTTACT 188
DB 1 CACGTTGTTCCGTATGACCTTGGAAAGCCCTAGTCCCAAGAGAGCCCTTGGAGAAATTTACT 60

QY 189 CCAACAAAGGCAACAGACCGTGAAATAGATGCCAATGTCTAGCCCAAGAAAGACAAGAA 248
DB 61 CCAACAAAGGCAACAGACCGTGAGATAGATGCCAATGTCTAGCCCAAGAAAGACAAGAA 120

QY 249 TGCTGGAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAGAAA 294
DB 121 TGCTGGAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAGAAA 166

RESULT 15
US-09-918-995-21013
Sequence 21013, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 21013
LENGTH: 512
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(512)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-21013

Query Match 6.8%; Score 155.8; DB 10; Length 512;
Best Local Similarity 95.8%; Pred. No. 2e-30;
Matches 160; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 128 GCACGTTGTTCCGTATGACCTTGGAAAGCCCTAGTCCCAAGAGAGCCCTTGGAGAAATTTACT 187
DB 157 GCACGTTGTTCCGTATGACCTTGGAAAGCCCTAGTCCCAAGAGAGCCCTTGGAGAAATTTACT 216

QY 188 TCCCAACAAAGGCAACAGACCGTGAAATAGATGCCAATGTCTAGCCCAAGAAAGACAAGAA 247
DB 217 TCCCAACAAAGGCAACAGACCGTGAGATAGATGCCAATGTCTAGCCCAAGAAAGACAAGAA 276

QY 248 GTGCTGGAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAGAAA 294
DB 277 GTGCTGGAATTTTCCCAAGCAGGAAAGAACTCAGGTGAGCAGAAA 323

Search completed: July 16, 2004, 23:09:17
Job time : 1006 secs
```

this Page Blank (uspto)

5	395	17.2	549	14	R56187
6	316.6	13.8	425	14	R55914
7	284.4	12.4	438	14	CB852037
8	281.6	12.2	518	28	AQ615825
9	279.8	12.2	291	14	F06636
10	245	10.6	713	12	BM974937
11	184	8.0	412	14	R48157
12	180	7.8	202	12	B1040245
13	157.4	6.8	673	12	B1821379
14	157.4	6.8	740	12	BM718889
15	157.4	6.8	769	14	CB988507
16	157.4	6.8	826	10	BE904559
17	157.4	6.8	848	12	B1758161
18	157.4	6.8	891	12	B1818778
19	157.4	6.8	1053	9	AL543736
20	157.4	6.8	1087	12	BM546209
21	157.4	6.8	1138	9	AL548042
22	157.4	6.8	1154	12	BM545093
23	157.4	6.8	1190	9	AL550935
24	157.4	6.8	1201	13	EX433515
25	155.8	6.8	1109	12	BM805113
26	155	6.7	659	14	CD370121
27	155	6.7	665	14	CA426258
28	155	6.7	715	14	CA309734
29	155	6.7	758	12	BM680754
30	155	6.7	758	12	BOQ09676
31	155	6.7	835	12	BI868856
32	155	6.7	839	12	BG118495
33	155	6.7	938	13	EX348929
34	155	6.7	959	13	EX348930
35	154	6.7	598	12	BM671893
36	154	6.7	603	9	AW043551
37	154	6.7	731	12	BM977229
38	154	6.7	837	9	AI039852
39	153.4	6.7	582	10	AW119009
40	153.4	6.7	646	13	EX098479
41	153.4	6.7	689	13	BU626469
42	153.4	6.7	756	14	CA431861
43	152.2	6.6	1201	9	AL533672
44	150.8	6.6	596	10	AW956848
45	150.8	6.6	808	9	AI521846

ALIGNMENTS

RESULT 1	AG061490	661 bp	DNA	linear	GSS 03-NOV-2001
LOCUS	Pan troglodytes DNA, clone: PTB-049121.R, genomic survey sequence.				
DEFINITION	AG061490				
ACCESSION	AG061490				
VERSION	AG061490.1 GI:16613292				
KEYWORDS	GSS.				
SOURCE	Pan troglodytes (chimpanzee)				
ORGANISM	Pan troglodytes				
REFERENCE	1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.				
AUTHORS	Fujiyama, A., Hattori, M., Toyoda, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.				
TITLE	BAC end sequences of Library PTB				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 661)				
AUTHORS	Fujiyama, A., Hattori, M., Toyoda, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimbes@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)				
COMMENT	Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of				

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 16, 2004, 15:06:54 ; Search time 5805 Seconds
(without alignments)
11836.844 Million cell updates/sec

Title: US-10-043-715-1_COPY_157700_160000

Perfect score: 2301

Sequence: 1 ctgaaatgagctcccaagt.....tagctccacggagagccct 2301

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :	EST:
1:	em_estba:*
2:	em_esthum:*
3:	em_estin:*
4:	em_estmu:*
5:	em_estov:*
6:	em_estpl:*
7:	em_estro:*
8:	em_htc:*
9:	gb_est1:*
10:	gb_est2:*
11:	gb_htc:*
12:	gb_est3:*
13:	gb_est4:*
14:	gb_est5:*
15:	em_estfun:*
16:	em_estom:*
17:	em_gss_hum:*
18:	em_gss_inv:*
19:	em_gss_pln:*
20:	em_gss_vrt:*
21:	em_gss_fun:*
22:	em_gss_mam:*
23:	em_gss_mus:*
24:	em_gss_pro:*
25:	em_gss_fod:*
26:	em_gss_pig:*
27:	em_gss_vrl:*
28:	gb_gss1:*
29:	gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	568.8	24.7	661	29	AG061490 Pan trogl
2	544.8	23.7	624	29	AG157710 Pan trogl
3	530.4	23.1	532	28	B48925 RPK11-4A24
4	469.2	20.4	474	9	AW016885 UI-B10p

```

clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI
FEATURES
Source
1..661
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-049121.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match 24.7%; Score 568.8; DB 29; Length 661;
Best Local Similarity 98.5%; Pred. No. 1.6e-84;
Matches 595; Conservative 0; Mismatches 7; Indels 2; Gaps 2;
QY 1420 GAGCTCAGCCAGGAGTCATGAACTACTTCTGAGTTATTACTACTGATTTGTGGGCA 1479
Db |||||
59 GAGCTCAGCCAGGAGTCATGAACTACTTCTGAGTTGTGTACTACTGATTTGTGGGCA 118
QY 1480 GCCTCAGCTATCGGTTCTTCCACACCTGCTTATGAGAGTATCCATATTTATGTCGCAGG 1539
Db |||||
119 GCCTCAGCTATCGGTTCTTCCACACCTGCTTATGAGAGTATCCATATTTATGTCGCAGG 178
QY 1540 CCAGTAATGCTCCACACAGATCAGTTCTGAACTAAGCTGGAATTTTATGGGTTTT 1599
Db |||||
179 CCAGTAATGCTCCACACAGATCAGTTCTGAACTAAGCTGGAATTTTATGGGTTTT 238
QY 1600 ATTATGCCAATTAATAATCAACATTAACAGTTCTTCCCTCTGATTTCTCTGTAAACA 1659
Db |||||
239 ATTATGCCAATTAATAATCAACATTAACAGTTCTTCCCTCTGATTTCTCTGTAAACA 298
QY 1660 TTAGGCTGCAAAAAAATAATCTTTTAAAAATAATGCCATAAAGTATTTGCTCTGG 1719
Db |||||
299 TCAGGCTGTC-AAAAAATAATCTTTTAAAAATAATGCCATAAAGTATTTGCTCTGG 357
QY 1720 GCCTACTGATGCTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 1779
Db |||||
358 GCCTACTGATGCTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 417
QY 1780 AGATGGCCATAACTATTAACAAACCTATGCTGAGTTCTCTCAAGGAGGTCACATAGTAT 1839
Db |||||
418 AGATGGCCATAACTATTAACAAACCTATGCTGAGTTCTCTCAAGGAGGTCACATAGTAT 477
QY 1840 GAAGTTGGATGGGCTACGGAAGAAACAGAACAACTCTAGTTTATTTAAACCTGTA 1899
Db |||||
478 GAAGTTGGATGGGCTACGGAAGAAACAGAACAACTCTAGTTTATTTAAACCTGTA 537
QY 1900 TTTACTGCCACTTCCCTTTAGA-CTTGACCATATGACCCCTCGCTCCCATTTCTAAGCAT 1958
Db |||||
538 TTTACTGCCACTTCCCTTTAGACCTGGCCATATGACCCCTCGCTCCCATTTCTAAGCAT 597
QY 1959 AGGGCAGGCTTTATTTTCAATGGAATAGATATCACTTGAGGTTTTATCAAGAGTT 2018
Db |||||
598 AGGGCAGGCTTTATTTTCAACCGGTAATAGATATCACTTGAGGTTCTATCAAGAGTT 657
QY 2019 GCGG 2022
Db |||||
658 GCGG 661
RESULT 2
AG157710/c
LOCUS
DEFINITION Pan troglodytes DNA, clone: RP43-022L11.TJ, genomic survey
sequence.
ACCESSION AG157710

```

```

VERSION AG157710.1 GI:16687388
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
BAC end sequences of Library RPCI-43
Unpublished
2 (bases 1 to 624)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpanzee@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library RPCI-43. This BAC
end was generated during the R&D process and may have higher chance
of clone tracking errors.
PRIMERS
Sequencing: TJ
LIBRARY
Vector : pBACe3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
FEATURES
Source
1..624
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="RP43-022L11.TJ"
/sex="male"
/cell_type="lymphocytes"
/clone_lib="RPCI-43 Chimpanzee Male BAC Library"
ORIGIN
Query Match 23.7%; Score 544.8; DB 29; Length 624;
Best Local Similarity 96.5%; Pred. No. 1.5e-80;
Matches 579; Conservative 0; Mismatches 17; Indels 4; Gaps 2;
QY 704 TTAGAAAACAGTAGCAGAGGAGATCTATGATCTCTATAGATTAAAGAGCAAGAAATC 763
Db |||||
620 TTAGAAACAGTAGCAGAGGAGATCTATGATCTCTATAGATTAAAGAGCAAGAAATC 561
QY 764 CCTCTTAATATTTCCATGAAGCTCTGGAATGCAAAACCGATGCTCTGTACTTTTAGCA 823
Db |||||
560 CCTCTTAATATTTCCAGAAAGCTCTGGAATGCAAAACCGATGCTCTGTACTTTTAGCA 501
QY 824 CATACATTTTCATCTACAGGTAGATTTCCCAACCAAAATATATCCAGAGATGCTTTGTC 883
Db |||||
500 CATACATTTTCATCTACAGGTAGATTTCCCAACCAAAATATATCCAGAGATGCTTTGTC 441
QY 884 ATTGGTTTATATACAGGCTTTGCCCTCTCTGAGTCAATGTATTTACCACTTTCCCTGAGAA 943
Db |||||
440 ATTGGTTTATATACAGGCTTTGCCCTCTCTGAGTCAATGTATTTACCACTTTCCCTGAGAA 381
QY 944 ATCGAATATCAATTTGGGAGCGGACATTTAGAAAAGAAATCAAGTGTATGATTAATC 1003
Db |||||
380 ATCGAATATCAATTTGGGAGCGGACATTTAGAAAAGAAATCAAGTGTATGATTAATC 321
QY 1004 AAATTTCTCAATAAGTTGCAAGTTATTCAGATGGCCAAAGGAAAAATAAAGTCATTAGATA 1063
Db |||||
320 AAATTTCTCAATAAGTTGCAAGTTATTCAGATGGCCAAAGGAAAAATAAAGTCATTAGATA 261
QY 1064 GGGTGTGTAGATTTAGAACATCTGTTTTTCAGGTTTATGG---TCTTTTTTTTTTTT 1120
Db |||||
260 GGGTGTGTAGATTTAGAACATCTGTTTTTCAGGTTTATGGTCTTTTTTTTTTTTTTT 201
QY 1121 TTTTTTAAATAGGAAAAATGTGTTTGTGTGAGAGCCCAATGTCTATCCAAAAAGCTCTC 1180

```

```

Db      200 TTTTAAATAGGGAATGTTTGGTCAGAGCCCAATGTCATTCCTCAAAAGCTCTCTC 141
QY      1181 TTTTCTGCTGATGCTGCTGGAGAGAGAGGATCTGATAGGCAACATCATAGA 1240
Db      140 TTTTCTGCTGATGCTGCTGGAGAGAGAGGATCTGATAGGCAACATCATAGA 82
QY      1241 GTTCTCTGAGCTGCTCTTTGGTGATGAACCTTCCAAATCCTTAAACCTTTTGGAAATTCAC 1300
Db      81 GTTCTCTGAGCTGCTCTTTGGTGATGAACCTTCCAAATCCTTAAACCTTTTGGAAACCGC 22

RESULT 3
B48925/c
LOCUS      B48925
DEFINITION RPCI11-4A24.TP RPCI-11 Homo sapiens genomic clone RPCI-11-4A24,
            genomic survey sequence.
ACCESSION  B48925
VERSION    B48925.1 GI:2601162
KEYWORDS   GSS.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 532)
AUTHORS    Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
            Venter,J.C.
REFERENCE   Use of BAC End Sequences for Sequence-Ready Map Building
            Unpublished (1997)
            Other GSSs: RPCI11-4A24-TV
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: SP6
            Class: BAC ends.

FEATURES             Location/Qualifiers
     source           1..532
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7501175"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-4A24"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="vector: pBACs3.6; Site 1: EcoRI; Site 2: EcoRI;
                     RPCI11 Human Male BAC Library"

ORIGIN
Query Match      23.1%; Score 530.4; DB 28; Length 532;
Best Local Similarity 99.8%; Pred. No. 4e-78;
Matches 531; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      757 AAGATCCCTTAAATATTTCCATGAGCTCTGGAATGCAACCGATGCTCTGCTACT 816
Db      532 AAGATCCCTTAAATATTTCCATGAGCTCTGGAATGCAACCGATGCTCTGCTACT 473
QY      817 TTTAGCACATACCATTTTCATCTACAGTAGATTTCCCAACCAAAATATATCCAGAGATGC 876
Db      472 TTTAGCACATACCATTTTCATCTACAGTAGATTTCCCAACCAAAATATATCCAGAGATGC 413
QY      877 CTTTGTCTATGGGTATATACAGCCCTTGGCTCTCTGAGTCAATGTAATTACCATCTTCC 936

```

```

Db      412 CTTTGTCTATGGGTATATACAGCCCTTGGCTCTCTGAGTCAATGTAATTACCATCTTCC 353
QY      937 CTGAGAAATCGAAATCATTTTGGGAGCGGACATTTAGAAAAAGAAATCAAGTGTCTATG 996
Db      352 CTGAGAAATCGAAATCATTTTGGGAGCGGACATTTAGAAAAAGAAATCAAGTGTCTATG 293
QY      997 GATAAATCAAAATCTTCAATAAGTTGAGTTTATTCAGATGCCCCAAGGAAAAATAAAGTCA 1056
Db      292 GATAAATCAAAATCTTCAATAAGTTGAGTTTATTCAGATGCCCCAAGGAAAAATAAAGTCA 233
QY      1057 TTAGATAGGGTTGGTAGAATTTAGAACATGCTGTTTTCAGGTTTATGGTCTTTTTTTTTT 1116
Db      232 TTAGATAGGGTTGGTAGAATTTAGAACATGCTGTTTTCAGGTTTATGGTCTTTTTTTTTT 173
QY      1117 TTTTCTTTTAAATAGGGAATGCTTTGGTGACAGAGCAATGTCATTCCAAAAAGCTC 1176
Db      172 TTTTCTTTTAAATAGGGAATGCTTTGGTGACAGAGCAATGTCATTCCAAAAAGCTC 113
QY      1177 TCTCTTTTCTGTCAGTCATGCTGCGGACAGAGAGGATCTGGATTAGGCAACATCA 1236
Db      112 TCTCTTTTCTGTCAGTCATGCTGCGGACAGAGAGGATCTGGATTAGGCAACATCA 53
QY      1237 TAGAGTTGCTCTGAGCTGCTCTTTGGTGATTAACCTTCCAAATCCTTAAACTT 1288
Db      52 TAGAGTTGCTCTGAGCTGCTCTTTGGTGATTAACCTTCCAAATCCTTAAACTT 1

RESULT 4
AW016885
LOCUS      AW016885
DEFINITION UT-H-B10p-abi-c-07-0-UI.s1 NCI_CGAP_Sub2 Homo sapiens cDNA clone
            IVAGS:2711820 3', mRNA sequence.
ACCESSION  AW016885
VERSION    AW016885.1 GI:5865642
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 474)
REFERENCE   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
            Unpublished (1997)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            The sequence contained an oligo-dT track that was present in the
            oligonucleotide that was used to prime the synthesis of first
            strand cDNA and therefore this may represent a bonafide poly. A
            tail. cDNA library preparation: M.B. Soares lab Clone distribution:
            NCI-CGAP clone distribution information can be found through the
            I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html The following repetitive
            elements were found in this cDNA sequence: 1-22,
            >AT rich/Low complexity
            Seq_primer: M13 Forward
            POLYA=Yes.

FEATURES             Location/Qualifiers
     source           1..474
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2711820"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_Sub2"
                     /note="vector: p7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site 1: Not 1; Site 2: Eco RI; The
                     NCI_CGAP Sub2 library is a subtracted library derived from
                     BI. BI constitutes a mixture of 21 normalized or
                     subtracted NCI CGAP libraries: NCI_CGAP Co4,
                     NCI_CGAP Pr22, NCI_CGAP Pr28, NCI_CGAP Co10,
                     NCI_CGAP Co16, NCI_CGAP_Kid5, NCI_CGAP_Kid12,

```


FEATURES
source

Location/Qualifiers
1. .438
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-FNO-aes-p-20-0-UI"
/tissue_type="Human Lung Epithelial cells"
/lab_host="PH10B (Life technologies) (T1 phage resistant)"
/clone_lib="UI-CF-FNO"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoRI; Site 2: Not I; UI-CF-FNO is a subtracted cDNA library derived from two normalized Human lung epithelial cell libraries (EN1 and DUN). The library was subtracted according to according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. For additional information, contact: bento-soares@uiowa.edu
TAG_SEQ=None found"

ORIGIN

Query Match 12.4%; Score 284.4; DB 14; Length 438;
Best Local Similarity 98.0%; Pred. No. 2.2e-37;
Matches 288; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTGAATGATGCTCCCAAGTCTATGATGAGCTCTTGTGTCGCCAGTGGATAGGTG 60
DB |||||
84 CTGAATGATGCTCCCAAGTCTATGATGAGCTCTTGTGTCGCCAGTGGATAGGTG 143
QY 61 TGTCCATGTCATTTTAAAGACTATTAAATACACTAATATAGTTCTTCTCTTTGG 120
DB |||||
144 TGTCCATGTCATTTTAAAGACTATTAAATACACTAATATAGTTCTTCTCTTTGG 203
QY 121 ATAATAGGACGTTGTTCCGATGAGCTTGAAGCCCTAGTCCCAAGAGCCCTTGAGA 180
DB |||||
204 ATAATAGGACGTTGTTCCGATGAGCTTGAAGCCCTAGTCCCAAGAGCCCTTGAGA 263
QY 181 ATTACTTCCCAAGGCAACAGACCGTGAATATAGATGCCATGTGTCAGCAAAAG 240
DB |||||
264 ATTACTTCCCAAGGCAACAGACCGTGAATATAGATGCCATGTGTCAGCAAAAG 323
QY 241 ACAAGAGTGTGGAATTTTGCACAGCAGAAAGAACTCAGTGAGCAGAA 294
DB |||||
324 ACAAGAGTGTGGAATTTTGCACAGCAGAAAGAACTCAGTGAGCAGAA 377

RESULT 8
AQ615825/c

LOCUS HS:5146_B2_C10_SP6E_RPCI-11 Human Male BAC Library Homo sapiens
DEFINITION HS:5146_B2_C10_SP6E_RPCI-11 Human Male BAC Library Homo sapiens
Genomic clone Plate=722 Col=20 Row=F, genomic survey sequence.
ACCESSION AQ615825
VERSION AQ615825.1 GI:5077101
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

TITLE

Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

JOURNAL
MEDLINE
PUBMED

Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (<http://www.htsc.washington.edu>). BAC end Web Server: <http://www.htsc.washington.edu>
Plate: 722 row: F column: 20
Seq primer: SP6
Class: BAC ends

High quality sequence stop: 518.

FEATURES
source

Location/Qualifiers
1. .518
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=722 Col=20 Row=F"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at EcoRI sites"

ORIGIN

Query Match 12.2%; Score 281.6; DB 28; Length 518;
Best Local Similarity 87.2%; Pred. No. 5.9e-37;
Matches 333; Conservative 0; Mismatches 44; Indels 5; Gaps 2;

QY 929 CACTTTCCCTGAGAAATCGAAATCATTTGGGGAGCGGACATTTAGAAAAGATCAAA 988
DB |||||
475 CCCCCTTCTCGAGAAATCAAAATCATTTGGGGGGCGGACATTTAGAAAAGATCAAA 416
QY 989 GT-GTCATGGAATCAAAATCTTCAATAGTTGCAGTTATTTCAGATGCGCAAGGAAAA 1047
DB |||||
415 GTGGTCTAGGAGATCCAAATCTTCAATAGTTGCAGTTATTTCAGATGCGCAAGGAAAA 356
QY 1048 ATAAAGTCTAGTAGAGGTTGGTAGAATTTAGAACATGCTGTTTTCAGGTTATGGTC 1107
DB |||||
355 ATAAAGTCTTAGTAGAGGTTGGTAGAATTTAGAACATGCTGTTTTCAGGTTATGGTC 296
QY 1108 T----TTTTTTTTTTTTTTTTTAAATAGGAAATGTGTTTGGTCAGAGCCCAATGTCA 1163
DB |||||
295 TTCTCTATTTTACTTTTATATATATACATAGAAATGTGTTTGGTCAGAGCCCAATGTCA 236
QY 1164 TTCCAAAAGCTCTCTTTTCTGGTCAGTCATGCTGGGACAGAGAGGATCTGGA 1223
DB |||||
235 TTCCAAAAGCTCTCTTTTCTGGTCAGTCATGCTGGGACAGAGAGGATCTGGA 176
QY 1224 TTAGCAACATCATAGAGTTGCTCTGAGCTGCTCTTTGGTGATACCCCTTCCAATCTTA 1283
DB |||||
175 TTAGCAACATCATAGAGTTGCTCTGAGCTGCTCTTTGGTGATACCCCTTCCAATCTTA 116
QY 1284 AACTTTTGGAATTCACAGCT 1305
DB |||||
115 CCTATTAGGAATAACAGAGTCT 94

RESULT 9
F06636

LOCUS F06636
DEFINITION HSC1F5031 normalized infant brain cDNA Homo sapiens linear EST 20-FEB-1995
c-1fe03, mRNA sequence.
ACCESSION F06636
VERSION F06636.1 GI:672236
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 291)
Auffray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes,M.D., Duprat,S., Houlgatte,R., Jumeau,M.N., Lamy,B.,

Lorenzo, F., Mitchell, H., Mariage-Samson, R., Pietu, G., Pouliot, Y.,
 Sebastiani-Kabakthchis, C. and Tessier, A.
 IMAGE: Molecular integration of the analysis of the human genome
 and its expression
 C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
 95277534
 MEDLINE
 PUBMED
 COMMENT

Contact: Genethon
 Genexpress-Genethon
 Genethon Centre de recherche sur le Genome Humain
 1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
 Tel: 33169472800
 Fax: 33160778698
 Email: genexpress@genethon.fr

Genexpress_library_id: C; Genexpress_sequence_id: ylc-1fe03
 Seq primer: (-21)M13-universal.
 Location/Qualifiers
 1. .291

FEATURES

source
 /organism="Homo sapiens"
 /mol_type="rRNA"
 /db_xref="taxon:9606"
 /clone="c-1fe03"
 /sex="Female"
 /tissue_type="total brain"
 /dev_stage="3 months old"
 /clone_lib="normalized infant brain cDNA"
 /note="Organ: brain; Vector: lafmid BA; Site 1: HindIII;
 Site 2: NotI; sex=Female; dev stage=3 months old;
 isolate=muscular atrophy patient; tissue_type=total
 brain; total mRNA was oligo-(dT) primed and directionally
 cloned 5' -> 3' into the HindIII -> NotI sites of the
 lafmid BA vector. Clone library from B.Souares, Psychiatry
 Dept. Columbia University, USA. Normalization method:
 Bento Soares, P.N.A.S in press"

ORIGIN

Query Match 12.2%; Score 279.8; DB 14; Length 291;
 Best Local Similarity 96.6%; Pred. No. 1.6e-36;
 Matches 281; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1565 TTCTGAACCTACCTGGAAATTTTATGGGTTTATTATGCAACTATTAATCAACAT 1624
 Db 1 TTNTGAACCTACCTGGAAATTTTATGGGTTTATTATGCAACTATTAATCAACAT 60

QY 1625 TACAGTTCTTCCCTCTGTATTTCTCTGTAAACATTAGGCTGCAAAAAAATAATCT 1684
 Db 61 TACAGTTCTTCCCTCTGTATTTCTCTGTAAACATTAGGCTGCAAAAAAATAATCT 120

QY 1685 TTATAAATAATTCCTAATAGTATTTCTCTGGCTCTGCTATGCTTCTTTCTTTT 1744
 Db 121 TTTTAAAAATAATNGCCATAAAGTATTTGCNCTGGGCTCTGCTATGCTTCTTTCTTTT 180

QY 1745 TCTCTCTTTTCAACTAAGTCACCGTCAATTTATTAGATGGCCATACTATTCAAACT 1804
 Db 181 NCTCTCTTTTCAACTAAGTCACCGTCAATTTATTAGATGGCCATACTATTCAAACT 240

QY 1805 ATGCTGAGTTCTCTAAGGAGGCTCATAGTATGATGAAGTTGGATGGG 1855
 Db 241 ATGCNAAGTTCTCTAAGGAGGCTCGCATAGTATGATGAAGTTGGATGGG 291

RESULT 10
 BM974937/c
 LOCUS
 DEFINITION
 UI-CF-EC1-acd-d-12-0-UI.s1 UI-CF-EC1 Homo sapiens cDNA clone
 UI-CF-EC1-acd-d-12-0-UI 3', mRNA sequence.
 BM974937
 ACCESSION
 VERSION
 BM974937.1 GI:19592528
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
 1 (bases 1 to 713)
 AUTHORS
 TITLE
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL
 Genome Res. 6 (9), 791-806 (1996)
 MEDLINE
 PUBMED
 COMMENT

Contact: McCray, PB
 McCray Lab
 University of Iowa
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
 Tel: 319 356 4866
 Fax: 319 356 7171
 Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com) or from Open Biosystems
 (www.openbiosystems.com).
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

source
 1. .713

/organism="Homo sapiens"
 /mol_type="rRNA"
 /db_xref="taxon:9606"
 /clone="UI-CF-EC1-acd-d-12-0-UI"
 /tissue_type="lung"
 /dev_stage="Adult and Fetal"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-CF-EC1"
 /notes="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
 modified polylinker; Site 1: EcoR I; Site 2: Not I;
 UI-CF-EC1 is a normalized cDNA library containing the
 following tissue(s): Normal lung from adult and from fetal
 day 64, day 87, week 19 and week 42. The library was
 constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pT73-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (dT)18 tail. The sequence tag for this library is
 AAGTGTCTTAC.

TAG TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
 and 380-383
 TAG LIB=UI-CF-EC1
 TAG_SEQ=AAGTGTCTTAC"

ORIGIN

Query Match 10.6%; Score 245; DB 12; Length 713;
 Best Local Similarity 98.2%; Pred. No. 5.5e-31;
 Matches 267; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 2031 GAAAGTTCACAAACAGATTCAGGTTTGTGTGCGCAGATTCATAATTTACATGTTCTT 2090
 Db 713 GAAAGTTCACAAACAGATTCAGG-TTGTGTGCGCAGATTCATAATTTACATGTTCTT 655

QY 2091 TTGCGAA-AGGGTGATTTTTTAAATAACATTTGTTTCTTCTTATCTTCGTTTATTAGG 2149
 Db 654 NTGCCAANAGGGTGATTTTTTAAATAACATTTGTTTCTTCTTATCTTCGTTTATTAGG 595

QY 2150 TCGGAGACCATGAGAAACAGCGTCAATCATCTTTTCATCATCCAGCTGGAAGCAAG 2209
 Db 594 TCGGAGACCATGAGAAACAGCGTCAATCATCTTTTCATCATCCAGCTGGAAGCAAG 535

QY 2210 CCCTCCAGAGCGCTTATGTGACCCACAAACCGAGCATTGGTGACAGACCTTCGGGGCC 2269

```

Db      534  CCTCCAGAGCGTTATGTGACCCACACGAGCACATTTGGTGACAGACCTTCGGGGCC 475
QY      2270  TGTCTGAAGCCATAGCCTCCACGGAGCCCT 2301
Db      474  TGTCTGAAGCCATAGCCTCCACGGAGCCCT 443

RESULT 11
R48197/c
LOCUS      R48197              412 bp      mRNA      linear      EST 18-MAY-1995
DEFINITION YJ63C09.r1 Soares breast 2NbhSt Homo sapiens cDNA clone
IMAGE:153424 5' similar to gb:S56805 ENDOTHELIN-1 PRECURSOR
(HUMAN);, mRNA sequence.
R48197
ACCESSION R48197.1 GI:810223
VERSION    EST.
KEYWORDS  Homo sapiens (human)
SOURCE    Homo sapiens
ORGANISM  Homo sapiens (human)
REFERENCE 1 (bases 1 to 412)
AUTHORS  Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevisan,S.E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
CONTACT: Wilton RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 849
High quality sequence stops: 274
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 849 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 274.
Location/Qualifiers
1..412
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:565741"
/db_xref="taxon:9606"
/clone="IMAGE:153424"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares Breast 2NbhSt"
/note="Organ: breast; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer
[5'
TGTTACCAATCTGAAGTGGAGCGCGCGCCCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of a modified pT7T3 vector (Pharmacia).
Library went through one round of normalization to a Cot =
230. Library constructed by Bento Soares and M.Fatima
Bonaldo."

ORIGIN
Query Match 8.0%; Score 184; DB 14; Length 412;
Best Local Similarity 100.0%; Pred. No. 9.4e-21;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CTGAAATGATCTCCCAAGTCTATGTGATGAGCTCTCTGTGTGCCAGTGAATAGGTG 60
Db      184  CTGAAATGATGCTCCCAAGTCTATGTGATGAGCTCTCTGTGTGCCAGTGAATAGGTG 125
QY      61  TGTCCATGTGTCAATTTTAAAGACTATTAAATTAACATAATAGTTTCTTCTCTCTTTG 120

Db      124  TGTCCTATGTCTCATTTTAAAGACTATTAAATTAACATAATAGTTTCTTCTCTCTTTGG 65
QY      121  ATAATAGGACGCTTGTTCGTTATGGACTTGGAGCCCTAGGTCCAGAGAGCCTTGGAGA 180
Db      64  ATAATAGGACGCTTGTTCGTTATGGACTTGGAGCCCTAGGTCCAGAGAGCCTTGGAGA 5
QY      181  ATTT 184
Db      4  ATTT 1

RESULT 12
BIO40245
LOCUS      BIO40245              202 bp      mRNA      linear      EST 14-JUN-2001
DEFINITION CM4-NT0287-010201-839-f02 NT0287 Homo sapiens cDNA, mRNA sequence.
ACCESSION BIO40245
VERSION    BIO40245.1 GI:14446871
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 202)
AUTHORS  Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
MEDLINE
10737800
PUBMED
COMMENT  Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=CM4&t2=CM4-NT0287-
010201-839-f02&t3=2001-02-01&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 17
High quality sequence stop: 202.
Location/Qualifiers
1..202
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="NT0287"
/note="Organ: nervous tumor; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORISTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match 7.8%; Score 180; DB 12; Length 202;
Best Local Similarity 94.9%; Pred. No. 6.5e-20;
Matches 186; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      1822  GCAGGGTCACATAGTATGATGAAGGTTGGATGGGCTACCGAAGAACCCAGAACACTCTA 1891
Db      7  GCCGGGCCGCATAGTATGATGAAGGTTGGATGGGCTACCGAAGAACCCAGAACACTCTA 66

```


QY 188 TCCCAAGGCAACAGACCGTGAAATAGATGCCAATGCTAGCCCAAAAGACAGAA 247
Db 546 TCCCAAGGCAACAGACCGTGAAATAGATGCCAATGCTAGCCCAAAAGACAGAA 605
QY 248 GTGCTGGAATTTTCCCAAGCAGGAGAAAGAACTCAGGTGAGCAGAAA 294
Db 606 GTGCTGGAATTTTCCCAAGCAGGAGAAAGAACTCAGGTGAGCAGAAA 652

RESULT 15
CB988507
LOCUS
DEFINITION CB988507 769 bp mRNA linear EST 01-MAY-2003
IMAGE:30340493 5', mRNA sequence.
ACCESSION CB988507
VERSION
KEYWORDS
SOURCE CB988507.1 GI:30283027
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 769)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDA370 row: h column: 06
High quality sequence stop: 620.
Location/Qualifiers
1..769

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30340493"
/tissue_type="Human Placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC 147"
/note="Organ: placenta; Vector: pBluescriptR; Site 1:
ali-Xhol; Site 2: BamH; Oligo-dT primed using primer
5'-TTTTTTTTTTTATVN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: This is
a NIH_MGC library."

ORIGIN

Query Match 6.8%; Score 157.4; DB 14; Length 769;
Best Local Similarity 96.4%; Pred. NO. 1.6e-16;
Matches 161; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 128 GCACGTTGTCGATGACCTGGAAGCCCTAGGTCCAAAGAGACGCTTGGAGATTTACT 187
Db 529 GCACGTTGTCGATGACCTGGAAGCCCTAGGTCCAAAGAGACGCTTGGAGATTTACT 598
QY 188 TCCCAAGGCAACAGACCGTGAAATAGATGCCAATGCTAGCCCAAAAGACAGAA 247
Db 589 TCCCAAGGCAACAGACCGTGAAATAGATGCCAATGCTAGCCCAAAAGACAGAA 648
QY 248 GTGCTGGAATTTTCCCAAGCAGGAGAAAGAACTCAGGTGAGCAGAAA 294
Db 649 GTGCTGGAATTTTCCCAAGCAGGAGAAAGAACTCAGGTGAGCAGAAA 695